

①9 RÉPUBLIQUE FRANÇAISE  
INSTITUT NATIONAL  
DE LA PROPRIÉTÉ INDUSTRIELLE  
PARIS

①1 N° de publication : **2 767 336**  
(à n'utiliser que pour les  
commandes de reproduction)  
②1 N° d'enregistrement national : **97 10404**

⑤1 Int Cl<sup>6</sup> : C 12 N 15/70, C 12 N 15/31, 1/21, C 12 Q 1/68, C 07 K  
14/35, 19/00, G 01 N 33/68, 33/569, A 61 K 39/04, 48/00 //  
(C 12 N 1/21, C 12 R 1:32)

⑫

## DEMANDE DE BREVET D'INVENTION

A1

②2 Date de dépôt : 14.08.97.

③0 Priorité :

④3 Date de mise à la disposition du public de la  
demande : 19.02.99 Bulletin 99/07.

⑤6 Liste des documents cités dans le rapport de  
recherche préliminaire : *Se reporter à la fin du  
présent fascicule*

⑥0 Références à d'autres documents nationaux  
apparentés :

⑦1 Demandeur(s) : INSTITUT PASTEUR — FR.

⑦2 Inventeur(s) : GICQUEL BRIGITTE, PORTNOI  
DENIS, LIM ENG MONG et PELICIC VLADIMIR.

⑦3 Titulaire(s) :

⑦4 Mandataire(s) : REGIMBEAU.

⑤4 VECTEURS RECOMBINANTS, POLYNUCLEOTIDE CODANT POUR UN POLYPEPTIDE PD428 DE 12KD DE  
MYCOBACTERIES APPARTENANT AU COMPLEXE DE MYCOBACTERIUM TUBERCULOSIS ET  
APPLICATIONS AU DIAGNOSTIC ET A LA PREVENTION DE LA TUBERCULOSE.

⑤7 L'invention a pour objet de nouveaux vecteurs recom-  
binants se répliquant chez les mycobactéries, des procédés  
de criblage incluant lesdits vecteurs, des polypeptides, dont  
le polypeptide DP428, et leur séquence nucléique, corres-  
pondant à des polypeptides exportés retrouvés dans les  
mycobactéries appartenant au complexe de Mycobacterium  
tuberculosis.

L'invention concerne également des procédés et des  
kits de détection de mycobactéries appartenant au complexe  
de Mycobacterium tuberculosis dans un échantillon bio-  
logique utilisant lesdits polypeptides, leurs anticorps  
spécifiques ou lesdits polynucléotides, ainsi que des com-  
positions immunogènes ou vaccinales pour la prévention et/  
ou le traitement d'infections provoquées par des mycobac-  
téries appartenant audit complexe, en particulier la tubercu-  
lose.

FR 2 767 336 - A1



Vecteurs recombinants, polynucléotide codant pour un polypeptide DP428 de 12kD de mycobactéries appartenant au complexe de *Mycobacterium tuberculosis* et applications au diagnostic et à la prévention de la tuberculose.

5

L'invention a pour objet de nouveaux vecteurs recombinants de criblage, de clonage et/ou d'expression se répliquant chez les mycobactéries. L'invention concerne également un  
10 polypeptide, dénommé DP428, d'environ 12kD correspondant à une protéine exportée retrouvée dans les mycobactéries appartenant au complexe de *Mycobacterium tuberculosis*. L'invention vise aussi un polynucléotide comprenant une séquence codant pour ce polypeptide. Elle concerne également l'utilisation du  
15 polypeptide ou de fragments de celui-ci et des polynucléotides codant pour ces derniers pour la réalisation de moyens de détection in vitro de la présence d'une mycobactérie appartenant au complexe de *Mycobacterium tuberculosis* dans un échantillon biologique. L'invention vise enfin l'utilisation du  
20 polypeptide ou de fragments de celui-ci ainsi que des polynucléotides codant pour ces derniers en tant que moyens destinés à la préparation d'une composition immunogène, susceptible d'induire une réponse immunitaire dirigée contre les mycobactéries appartenant au complexe de *Mycobacterium*  
25 *tuberculosis*, ou d'une composition vaccinale pour la prévention et/ou le traitement d'infections provoquées par des mycobactéries appartenant audit complexe, en particulier la tuberculose.

30

Le genre *Mycobacterium*, qui comprend au moins 56 espèces différentes, inclut des pathogènes humains majeurs tels que *M. leprae* et *M. tuberculosis*, les agents responsables de la lèpre et de la tuberculose, qui restent des problèmes graves de santé  
35 publique dans le monde entier.

La tuberculose continue d'être un problème de santé publique dans le monde. Aujourd'hui, cette maladie est la cause de 2 à 3 millions de morts dans le monde et environ 8 millions de nouveaux cas sont observés chaque année (Bouvet, 1994). Dans  
5 les pays développés *M. tuberculosis* est la cause la plus commune des infections mycobactériennes. En France il apparaît environ 10 000 nouveaux cas par an et parmi les maladies à déclaration obligatoire c'est la tuberculose qui comprend le plus grand nombre de cas. La vaccination par le BCG (Bacille de  
10 Calmette et Guérin), une souche avirulente dérivée de *M. bovis* et qui est très utilisé comme vaccin contre la tuberculose, est loin d'être efficace au sein de toutes les populations. Cette efficacité varie environ de 80% dans les pays occidentaux comme l'Angleterre, à 0% en Inde (résultats du dernier essai de  
15 vaccination à Chingleput., publiés en 1972 dans Indian J. Med. Res.). De plus, l'apparition de souches de *M. tuberculosis* résistantes aux antituberculeux et le risque accru chez les patients immunodéprimés, patients atteints du SIDA, de développer une tuberculose, rendent nécessaire la mise au point  
20 de méthodes rapides, spécifiques et fiables pour le diagnostic de la tuberculose. Par exemple, une étude épidémiologique réalisée en Floride, et dont les résultats ont été publiés en 1993 dans AIDS thérapies, a montré que 10% des malades atteints de SIDA sont atteints de tuberculose au moment du diagnostic du  
25 SIDA ou 18 mois avant celui-ci. Chez ces malades, la tuberculose apparaît dans 60% des cas sous une forme disséminée donc non repérable par les critères de diagnostic classiques comme la radiographie pulmonaire ou l'analyse de crachats.

Actuellement, une certitude sur le diagnostic apporté par  
30 la mise en évidence de bacilles cultivables dans un prélèvement provenant du malade n'est obtenue que pour moins de la moitié des cas de tuberculose, même dans les cas de tuberculose pulmonaire. Le diagnostic de la tuberculose et des autres mycobactéries apparentées est donc difficile à réaliser, et  
35 cela pour différentes raisons : les mycobactéries sont souvent présentes en faible quantité, leur temps de génération est très

long (24h pour *M. tuberculosis*) et leur culture est difficile. (Bates et al., 1986).

D'autres techniques sont utilisables en clinique, pour identifier une infection mycobactérienne :

5 a). L'identification directe des microorganismes au microscope; cette technique est rapide, mais ne permet pas l'identification de l'espèce mycobactérienne observée et manque de sensibilité (Bates, 1979).

10 Les cultures, lorsqu'elles sont positives, ont une spécificité approchant 100% et permettent l'identification de l'espèce mycobactérienne isolée; néanmoins, comme précisé ci-dessus, la croissance des mycobactéries *in vitro* est longue (ne peut être réalisée qu'en 3 à 6 semaines de cultures répétées (Bates, 1979; Bates et al., 1986)) et coûteuse.

15 b). Les techniques sérologiques peuvent s'avérer utiles dans certaines conditions, mais leur utilisation est parfois limitée par leur sensibilité et/ou leur spécificité faibles (Daniel et al., 1987).

20 c). La présence de mycobactéries au sein d'un échantillon biologique peut aussi être déterminée par hybridation moléculaire avec de l'ADN ou de l'ARN en utilisant des sondes d'oligonucléotides spécifiques des séquences recherchées (Kiehn et al., 1987; Roberts et al., 1987; Drake et al., 1987). Plusieurs études ont montré l'intérêt de cette  
25 technique pour le diagnostic des infections à mycobactéries. Les sondes utilisées sont constituées d'ADN, d'ARN ribosomique ou de fragments d'ADN mycobactériens non caractérisés et provenant de banque de gènes. Le principe de ces techniques repose sur le polymorphisme des séquences nucléotidiques des  
30 fragments utilisés ou sur le polymorphisme des régions avoisinantes. Dans tous les cas, elles nécessitent l'utilisation de cultures et ne sont pas applicables directement sur les échantillons biologiques.

35 La faible quantité de mycobactéries présentes au sein d'un échantillon biologique et en conséquence la quantité faible d'ADN cible à détecter dans cet échantillon peut nécessiter le recours à une amplification spécifique *in vitro* de l'ADN cible



avant sa détection à l'aide de la sonde nucléotidique et en utilisant des techniques d'amplification *in vitro* telles que la PCR (amplification en chaîne à la polymérase

5 L'amplification spécifique de l'ADN par la technique PCR peut constituer la première étape d'un procédé de détection de la présence d'un ADN mycobactérien dans un échantillon biologique, la détection proprement dite de l'ADN amplifié étant effectuée dans un second temps à l'aide d'une sonde oligonucléotidique capable de s'hybrider spécifiquement à l'ADN  
10 amplifié.

Un test de détection de mycobactéries appartenant au complexe de *Mycobacterium tuberculosis*, par hybridation sandwich (test utilisant une sonde de capture et une sonde de détection) a été décrit par Chevrier et al. en 1993. le  
15 complexe de *Mycobacterium tuberculosis* est un groupe de mycobactéries qui comprend *M. bovis*-BCG, *M. bovis*, *M. tuberculosis*, *M. africanum* et *M. microti*.

Un procédé de détection de faibles quantités de mycobactéries, appartenant au complexe tuberculosis, par  
20 amplification génique et hybridation directement sur des échantillons biologiques a été mis au point. Ledit procédé utilise la séquence d'insertion IS6110 (Brevet européen EP 0 490 951 B1). Thierry et al. ont décrit en 1990 une séquence spécifique du complexe *Mycobacterium tuberculosis* et nommée IS  
25 6110. Certains auteurs ont proposé d'amplifier spécifiquement l'ADN provenant de *Mycobacterium* en utilisant des amorces nucléiques dans une méthode d'amplification, telle que la réaction de polymérase en chaîne (PCR). Patel et al. ont décrit en 1990 l'utilisation de plusieurs amorces nucléiques choisies  
30 à partir d'une séquence connue en tant que sonde dans l'identification de *M. tuberculosis*. Cependant, la longueur des fragments obtenue en utilisant ces amorces était différente de la longueur théorique attendue et plusieurs fragments de taille variable étaient obtenus. De plus, les auteurs ont observé  
35 l'absence d'hybridation des produits amplifiés avec le plasmide

ayant servi à déterminer les amorces. Ces résultats indiquent que ces amorces ne seraient pas appropriées dans la détection de la présence de *M. tuberculosis* dans un échantillon biologique et confirment la nature critique du choix des amorces. La même année, J.L. Guesdon et D. Thierry ont décrit une méthode de détection de *M. tuberculosis*, de grande sensibilité, par amplification d'un fragment d'ADN de *M. tuberculosis* localisé au sein de la séquence IS6110 (Brevet européen EP 461 045) à l'aide d'amorces générant des fragments d'ADN amplifié de longueur constante, même lorsque le choix des amorces conduisait à l'amplification de fragments longs (de l'ordre de 1000 à 1500 bases) où le risque d'interruption de la polymérisation est élevée en raison des effets de la structure secondaire de la séquence. D'autres amorces spécifiques de la séquence IS6110 sont décrites dans le brevet européen N° EP-0490 951.

Les inventeurs ont montré (résultats non publiés) que certains isolats cliniques de *Mycobacterium tuberculosis* étaient exempts de la séquence d'insertion IS6110 et ne pouvaient donc être détectés à l'aide des oligonucléotides spécifiques de cette séquence pouvant conduire ainsi à des résultats de diagnostic faussement négatifs. Ces résultats confirment une observation similaire faite par Yuen et al. en 1993. L'impossibilité de détecter ces souches pathogènes potentiellement présentes dans un échantillon biologique prélevé sur un patient est ainsi susceptible de conduire à des difficultés voire des erreurs de diagnostic.

*M. bovis* et *M. tuberculosis*, les agents causals de la tuberculose, sont des bactéries facultatives intracellulaires. Ces agents ont développé des mécanismes pour assurer leur survie et leur répllication à l'intérieur du macrophage, un des types cellulaires qui est supposé éradiquer l'invasion par des microorganismes. Ces agents sont capables de moduler l'évolution normale de leur phagosome et de les empêcher de se différencier en un compartiment acide riche en hydrolase (5, 6,

23, 26). Cependant, cette modulation n'est possible que si la bactérie est vivante au sein du phagosome, suggérant que des composés synthétisés de manière active et/ou sécrétés à l'intérieur de la cellule font partie de ce mécanisme. Des protéines exportées sont probablement impliquées dans ce mécanisme. En dépit des problèmes majeurs de santé liés à ces organismes pathogènes, on sait peu de choses sur leurs protéines exportées et/ou sécrétées. Des analyses en SDS-PAGE de filtrat de culture de *M. tuberculosis* montrent au moins 30 protéines sécrétées (1,19,38). Certaines d'entre elles ont été caractérisées, leurs gènes clonés et séquencés (7,35,37). D'autres, bien qu'il s'agisse d'antigènes immunodominants d'importance majeure pour induire une immunité protectrice (2,21), ne sont pas totalement identifiés. En outre, il est probable que de nombreuses protéines exportées restent fixées sur la membrane cellulaire et par conséquent ne soient pas présentes dans les surnageants de culture. Il a été montré que les protéines localisées à la surface externe de diverses bactéries pathogènes, telles que l'invasine de 103 kDa de *Yersinia Pseudotuberculosis* (14) ou l'internaline de 80 kDa de *Listeria monocytogenes* (10) jouent un rôle important dans les interactions avec les cellules hôtes et par conséquent, dans la pathogénicité comme dans l'induction de réponses protectrices. Ainsi, une protéine liée à la membrane pourrait être importante pour l'infection à *M. tuberculosis* comme pour l'induction de réponse protectrice contre cette infection. Ces protéines pourraient revêtir un intérêt certain pour la préparation de vaccins.

Récemment, il a été décrit l'adaptation aux mycobactéries d'une méthodologie génétique pour l'identification et la sélection phénotypique de protéines exportées (15). Cette méthode utilise la phosphatase alcaline (PhoA) périplasmique d'*E. coli*. Un vecteur plasmidique a été construit permettant la fusion de gènes entre un gène *PhoA* tronquée et des gènes codant pour des protéines exportées (3, 11).

Par cette méthode, il a pu être identifier un gène de *M. tuberculosis* (erp (2)) présentant des homologies avec une protéine exportée de 28 kDa de *M. leprae*, qui est une cible fréquente des réponses humores de la forme lépromateuse de la lèpre, et avec une protéine présentant des motifs aminoacides caractéristiques de la désaturase de plante (*des*).

Cependant, cette méthode génétique d'identification de protéines exportées ne permet pas d'évaluer facilement l'expression intracellulaire des gènes correspondants. Une telle évaluation est d'une importance primordiale à la fois pour la sélection de bons candidats vaccins et pour la compréhension des interactions entre les bactéries et leurs cellules hôtes. L'induction de l'expression de facteur de virulence par contact de cellule cible pathogène a été décrite. C'est le cas par exemple pour le facteur de virulence Yops (18) de *Yersinia pseudotuberculosis*. *Shigella* par contact avec les cellules cibles relargue les protéines Ipa dans le milieu de culture, et *Salmonella* synthétise de nouvelles structures de surface.

Compte tenu de ce qui précède, il existe aujourd'hui un grand besoin de développer de nouveaux vaccins contre les mycobactéries pathogène ainsi que de nouveaux tests de diagnostic spécifiques, fiables et rapides. Ces développements nécessitent la mise au point d'outils spécifiques encore plus performants permettant, d'une part, d'isoler ou d'obtenir des séquences de nouveaux polypeptides spécifiques, notamment immunogènes, et, d'autre part, de mieux comprendre le mécanisme des interactions entre les bactéries et leurs cellules hôtes comme notamment l'induction de l'expression de facteur de virulence. Ceci est précisément l'objet de la présente invention.

Les inventeurs ont défini et réalisé dans ce but de nouveaux vecteurs permettant le criblage, le clonage et/ou l'expression de séquences d'ADN de mycobactéries afin d'identifier parmi ces séquences, des acides nucléiques codant

pour des protéines d'intérêt, de préférence des protéines exportées, pouvant être localisées sur la membrane bactérienne et/ou sécrètes, et d'identifier parmi ces séquences celles qui sont induites ou réprimées lors de l'infection.

5

L'invention vise aussi de nouveaux polypeptides et de nouveaux polynucléotides de mycobactéries ayant pu être isolés au moyen des vecteurs précédents et susceptibles d'entrer dans la réalisation de compositions pour la détection d'une  
10 infection par des mycobactéries, ou pour la protection contre une infection due à des mycobatéries.

L'invention a donc pour objet un vecteur recombinant de criblage, de clonage et/ou d'expression, caractérisé en ce  
15 qu'il se réplique chez des mycobactéries et en ce qu'il contient :

- 1) un réplicon fonctionnel chez les mycobactéries ;
- 2) un marqueur de sélection ;
- 3) une cassette reporteur comprenant :  
20 a) un site de clonage multiple (polylinker),  
b) éventuellement un terminateur de transcription actif chez les mycobactéries, en amont du polylinker,  
c) une séquence nucléotidique codante issue d'un gène codant pour un marqueur d'expression, d'exportation et/ou de  
25 sécrétion de protéine, ladite séquence nucléotidique étant dépourvue de son codon d'initiation et de ses séquences de régulation, et  
d) une séquence nucléotidique codante issue d'un gène codant pour un marqueur d'activité de promoteurs contenus dans  
30 le même fragment, ladite séquence nucléotidique étant pourvue de son codon d'initiation.

Le marqueur d'exportation et/ou de sécrétion est une séquence de nucléotides dont l'expression suivie de  
35 l'exportation et/ou de la sécrétion dépend des éléments de régulation qui contrôlent son expression.

Par ``séquences ou éléments de régulation de l'expression

de la production de polypeptides et de sa localisation'', on entend une séquence promotrice de la transcription, une séquence comprenant le site de liaison au ribosome (RBS), les séquences responsables de l'exportation et/ou la sécrétion  
5 telles que la séquence dite séquence signale.

Un premier marqueur intéressant d'exportation et/ou d'expression est une séquence codante issue du gène *phoA*. Le cas échéant, elle est tronquée de telle façon que l'activité phosphatase alcaline est cependant susceptible d'être restaurée  
10 lorsque la séquence codante tronquée est placée sous le contrôle d'un promoteur et d'éléments de régulation appropriés.

D'autres marqueurs d'exposition, d'exportation et/ou de sécrétion peuvent être utilisés. On citera à titre d'exemples une séquence du gène  $\beta$ -agarase, de la nucléase d'un  
15 staphylocoque ou d'une  $\beta$ -lactamase.

Parmi les marqueurs intéressants d'activité de promoteurs contenus dans le même fragment, on préfère une séquence codante issue du gène *luc* de luciférase de luciole pourvue de son codon d'initiation.

20 D'autres marqueurs d'activité de promoteurs contenus dans le même fragment peuvent être utilisés. On citera à titre d'exemples une séquence du gène de la GFP (Green Fluorescent Protein).

Le terminateur de transcription doit être fonctionnel chez  
25 les mycobactéries. Un terminateur avantageux est à cet égard le terminateur du coliphage T4 (tT4). D'autres terminateurs appropriés pour la réalisation de l'invention peuvent être isolés en utilisant la technique présentée dans les exemples, par exemple au moyen du vecteur cassette  $\Omega$ .

30 Un vecteur particulièrement préféré pour la réalisation de l'invention est un plasmide choisi parmi les plasmides suivants déposés à la CNCM (Collection Nationale de Cultures de Microorganismes, Paris, France) :

a) pJEVDA déposé à la CNCM sous le N° I-1797, le 12/12  
35 1996,

b) pJEVDb déposé à la CNCM sous le N° I-1906, le 25

juillet 1997,

c) pJEVDc déposé à la CNCM sous le N° I-1799 , le 12/12 1996,

d) pJEVD/*M. tuberculosis* déposé à la CNCM sous le N° I-  
5 1907, le 25 juillet 1997.

Pour la sélection, ou l'identification de séquences d'acides nucléiques de mycobactéries codant pour des polypeptides susceptibles d'être incorporés dans des  
10 compositions immunogènes, ou antigéniques pour la détection d'une infection, ou susceptibles d'induire ou de réprimer un facteur de virulence de mycobactéries, le vecteur de l'invention comprendra, en l'un des sites de clonage multiple du polylinker, une séquence de nucléotides d'une mycobactérie  
15 chez laquelle on détecte la présence de séquences correspondant à des polypeptides exportés et/ou sécrétés pouvant être induits ou réprimés lors de l'infection, ou encore exprimés ou produits de façon constitutive, leurs séquences promotrices et/ou régulatrices associées susceptibles de permettre ou de  
20 favoriser l'exportation et/ou la sécrétion desdits polypeptides d'intérêt, ou tout ou partie de gènes d'intérêt codant pour lesdits polypeptides.

De préférence, cette séquence est obtenue par fragmentation physique ou par digestion enzymatique de l'ADN  
25 génomique ou de l'ADN complémentaire d'un ARN d'une mycobactérie et de préférence d'une mycobactérie pathogène.

Selon un premier mode de réalisation de l'invention la digestion enzymatique de l'ADN génomique ou de l'ADN complémentaire est effectuée à partir de *M. tuberculosis*.

30 De préférence cet ADN est digéré avec une enzyme telle que sau3A, BclI, BglII.

D'autres enzymes de digestion telles que ScaI, ApaI, SacII, KpnI ou encore des nucléases ou des polymérases, peuvent naturellement être mises en oeuvre, dès lors qu'elles  
35 permettent l'obtention de fragments dont les extrémités peuvent être insérées dans l'un des sites de clonage du polylinker du vecteur de l'invention.

Le cas échéant, des digestions avec différentes enzymes seront effectuées simultanément.

Des vecteurs recombinants préférés pour la réalisation de  
5 l'invention sont choisis parmi les vecteurs recombinants suivants déposés à la CNCM :

- a) p6D7 déposé le 28 janvier 1997 à la CNCM sous le N°I-1814,
- b) p5A3 déposé le 28 janvier 1997 à la CNCM sous le N°I-1815,
- 10 c) p5F6 déposé le 28 janvier 1997 à la CNCM sous le N°I-1816,
- d) p2A29 déposé le 28 janvier 1997 à la CNCM sous le N°I-1817,
- e) pDP428 déposé le 28 janvier 1997 à la CNCM sous le N°I-1818,
- f) p5B5 déposé le 28 janvier 1997 à la CNCM sous le N°I-1819,
- 15 g) p1C7 déposé le 28 janvier 1997 à la CNCM sous le N°I-1820,
- h) p2D7 déposé le 28 janvier 1997 à la CNCM sous le N°I-1821,
- i) p1B7 déposé le 31 janvier 1997 à la CNCM sous le N°I-1843.

Parmi les plus préférés, on préfère le vecteur recombinant  
20 pDP428 déposé le 28 janvier 1997 à la CNCM sous le N°I-1818,

Les vecteurs de l'invention peuvent également être utilisés pour déterminer la présence de séquences d'intérêt, de préférence correspondant à des protéines exportées et/ou sécrétées, et/ou capables d'être induites ou réprimées lors de  
25 l'infection, notamment lors de la phagocytose par les macrophages, pJEVD et selon ce qui a été exposé précédemment, chez des mycobactéries telles que *M. africanum*, *M. bovis*, *M. avium* ou *M. leprae* dont on aura traité l'ADN ou l'ADNc avec des enzymes déterminées.

30 L'invention a également pour objet un procédé de criblage de séquences de nucléotides issues de mycobactéries pour déterminer la présence de séquences correspondant à des polypeptides exportés et/ou sécrétés pouvant être induits ou réprimés lors de l'infection, leurs séquences promotrices et/ou  
35 régulatrices associées susceptibles notamment de permettre ou de favoriser l'exportation et/ou la sécrétion desdits polypeptides d'intérêt, ou tout ou partie de gènes d'intérêt



codant pour lesdits polypeptides, caractérisé en ce qu'il met en oeuvre un vecteur recombinant selon l'invention.

L'invention concerne aussi un procédé de criblage, selon l'invention, caractérisé en ce qu'il comprend les étapes  
5 suivantes :

a) la digestion des séquences d'ADN de mycobactéries par au moins une enzyme déterminée et la récupération des fragments de digestion obtenus ;

b) l'insertion des fragments de digestion dans un site de  
10 clonage compatible avec l'enzyme de l'étape a) du polylinker d'un vecteur selon l'invention ;

c) si besoin, l'amplification du fragment de digestion contenu dans le vecteur, par exemple par réplication de ce dernier après insertion du vecteur ainsi modifié dans une  
15 cellule déterminée, par exemple *E coli* ;

d) la transformation des cellules hôtes par le vecteur amplifié à l'étape c), ou en l'absence d'amplification, par le vecteur de l'étape b) ;

e) la culture des cellules hôtes transformées dans un  
20 milieu permettant la mise en évidence du marqueur d'exportation et/ou de sécrétion, et /ou du marqueur d'activité de promoteurs contenu dans le vecteur ;

f) la détection des cellules hôtes positives (colonies positives) pour l'expression du marqueur  
25 d'exportation et/ou de sécrétion, et /ou du marqueur d'activité de promoteurs ;

g) l'isolement de l'ADN des colonies positives et l'insertion de cet ADN dans une cellule identique à celle de l'étape c) ;

h) la sélection des insertions contenues dans le vecteur, permettant l'obtention de clones positifs pour le marqueur  
30 d'exportation et/ou de sécrétion, et /ou pour le marqueur d'activité de promoteurs ;

i) l'isolement et la caractérisation des fragments d'ADN  
35 de mycobactéries contenues dans ces insérats.

La mise en oeuvre de ce procédé permet la construction de banques d'ADN comportant des séquences correspondant à des

polypeptides susceptibles d'être exportés et/ou sécrétés, et/ou susceptibles d'être induits ou réprimés lors de l'infection lorsqu'ils sont produits au sein de mycobactéries recombinantes. L'étape i) du procédé peut comprendre une étape  
5 de séquençage des insertions sélectionnées.

De préférence, dans le procédé selon l'invention, le vecteur utilisé est choisi parmi les plasmides pJEVDa (CNCM, N° I-1797), pJEVDb (CNCM, N° I-1906), pJEVDC (CNCM, N° I-1799) ou pJEVD/M. tuberculosis (CNCM, N°I-1907, 1907), et la  
10 digestion des séquences d'ADN de mycobactéries est effectuée au moyen de l'enzyme sau3A.

Selon un mode de réalisation préféré de l'invention, le procédé de criblage est caractérisé en ce que les séquences de mycobactéries sont issues d'une mycobactérie pathogène, par  
15 exemple de *M. tuberculosis*, *M. bovis*, *M. avium*, *M. africanum* ou *M. leprae*.

L'invention comprend également une banque d'ADN génomique ou d'ADNc complémentaire d'ARNm de mycobactérie, caractérisée en ce qu'elle est obtenue par un procédé  
20 comprenant les étapes a), b) et c) du procédé précédent selon l'invention, de préférence une banque d'ADN génomique ou d'ADNc complémentaire d'ARNm de mycobactéries pathogènes, de préférence de mycobactéries appartenant au groupe du complexe *Mycobacterium tuberculosis*, de préférence de *Mycobacterium*  
25 *tuberculosis*.

L'invention a également pour objet les séquences nucléotidiques de mycobactéries sélectionnées après la réalisation du procédé selon l'invention ci-dessus décrit.

Selon un mode de réalisation particulier de l'invention,  
30 des séquences préférées sont par exemple les fragments d'ADN de mycobactéries de séquence SEQ ID N°3, SEQ ID N°4, SEQ ID N°5, SEQ ID N°6, SEQ ID N°7, SEQ ID N°8, SEQ ID N°9 ou SEQ ID N°10, contenus respectivement dans les vecteurs p6D7 (CNCM, N°I-1814), p5A3 (CNCM, N°I-1815), p5F6 (CNCM, N°I-1816), p2A29  
35 (CNCM, N°I-1817), p5B5 (CNCM, N°I-1819), p1C7 (CNCM, N°I-1820), p2D7 (CNCM, N°I-1821) et p1B7 (CNCM, N°I-1843).

On préfère également les séquences nucléiques SEQ ID N° 11 et SEQ ID N°24.

Lorsque la séquence codante issue du gène marqueur  
5 d'exportation et/ou de sécrétion est une séquence issue du gène *phoA*, l'exportation et/ou la sécrétion du produit du gène *phoA*, le cas échéant tronqué, n'est obtenue que lorsque cette séquence est insérée en phase avec la séquence ou élément de régulation de l'expression de la production de polynucléotides  
10 et sa localisation placée en amont, qui contient les éléments contrôlant l'expression, l'exportation et/ou la sécrétion issus de séquence de mycobactéries.

Les vecteurs recombinants de l'invention peuvent bien  
entendu comprendre des sites de clonage multiples décalés de un  
15 ou deux nucléotides par rapport à un vecteur selon l'invention, permettant ainsi d'exprimer le polypeptide correspondant au fragment d'ADN de mycobactérie inséré et susceptible d'être traduit selon l'un des trois cadres de lecture possibles.

20 Par exemple les vecteurs préférés pJVEDb et pJVEDc de l'invention se distinguent du vecteur préféré pJVEDA par un décalage respectif de un et de deux nucléotides au niveau du site de clonage multiple.

Ainsi, les vecteurs de l'invention sont capables  
25 d'exprimer chacun des polypeptides susceptibles d'être codés par fragment d'ADN de mycobactérie inséré. Cesdits polypeptides, caractérisés en ce qu'ils sont susceptibles d'être codés par ledit fragment d'ADN et en ce qu'ils sont susceptibles d'être exportés et/ou sécrétés, et/ou induits ou  
30 réprimés lors de l'infection, font partie de l'invention.

L'invention a aussi pour objet des mycobactéries recombinantes contenant un vecteur recombinant selon l'invention décrit précédemment. Une mycobactérie préférée est une mycobactérie du type *M. smegmatis*.

35 *M. smegmatis* permet avantageusement de tester l'efficacité de séquences de mycobactéries, pour le contrôle de l'expression, de l'exportation et/ou de la sécrétion, et/ou de

l'activité de promoteurs d'une séquence donnée, par exemple d'une séquence codant pour un marqueur tel que la phosphatase alcaline et/ou la luciférase.

Une autre mycobactérie préférée est une mycobactérie du type *M. bovis*, par exemple la souche BCG utilisée actuellement pour la vaccination contre la tuberculose.

Une autre mycobactérie préférée est une souche de *M. tuberculosis*, *M. bovis* ou *M. africanum* possédant potentiellement tous les systèmes de régulation appropriés.

Les inventeurs ont ainsi caractérisé un polynucléotide constitué par une séquence de nucléotides présente chez toutes les souches testées de mycobactéries appartenant au complexe de *Mycobacterium tuberculosis*. Ce polypeptide, dénommé DP428 contient un cadre ouvert de lecture (ORF) codant pour un polypeptide d'environ 12 kD. Ce poids moléculaire (PM) correspond au PM théorique de la protéine mature obtenue après clivage de la séquence signale, le PM de la protéine ou polypeptide DP428 étant d'environ 10 kD après ancrage potentiel au peptidoglycane et coupure potentielle entre S et G du motif LPISG.

Ce polynucléotide inclut, d'une part, un cadre ouvert de lecture correspondant à un gène de structure et, d'autre part, les signaux de régulation de l'expression de la séquence codante en amont et en aval de cette dernière. Le polypeptide DP428 est composé d'un peptide signal, d'une région centrale hydrophile et d'une région C-terminale hydrophobe. Cette dernière se termine par deux résidus arginines (R), signal de rétention, et est précédé par un motif LPISG qui rappelle le motif LPXTG d'ancrage au peptidoglycane (Schneewind et al).

Par gène de structure aux fins de la présente invention, on entend un polynucléotide codant pour une protéine, un polypeptide ou encore un fragment de ces derniers, ledit polynucléotide ne comprenant que la séquence correspondant au cadre ouvert de lecture (ORF), ce qui exclut les séquences du côté 5' du cadre ouvert de lecture (ORF) qui dirigent l'initiation de la transcription.

Ainsi, l'invention concerne un polynucléotide de séquence SEQ ID N°1.

- Plus particulièrement, l'invention concerne un polynucléotide caractérisé en ce qu'il comprend un polynucléotide choisi parmi :
- a) un polynucléotide de séquence SEQ ID N°1 ou SEQ ID N°2,
  - b) un polynucléotide dont la séquence est complémentaire de la séquence du polynucléotide défini en a),
  - c) un polynucléotide dont la séquence comporte au moins 80% d'identité avec un polynucléotide défini en a) ou b),
  - d) un polynucléotide hybridant dans des conditions de forte stringence avec une séquence de polynucléotide défini en a), b) ou c),
  - e) un fragment de polynucléotide défini en a), b), c) ou d),
- et comportant au moins 8 nucléotides.

On entend par séquence nucléotidique, polynucléotide ou acide nucléique, selon la présente invention, aussi bien un ADN double brin, un ADN simple brin que des produits de transcription desdits ADN.

- Par polynucléotide de séquence complémentaire, on entend tout ADN dont les nucléotides sont complémentaires de ceux de la SEQ ID N°1 ou d'une partie de la SEQ ID N°1, et dont l'orientation est inversée.

- Par pourcentage d'homologie au sens de la présente invention, on entend un pourcentage d'identité entre les bases des deux polynucléotides homologues, ce pourcentage étant purement statistique et les différences entre les deux polynucléotides étant réparties au hasard et sur toute leur longueur.

- Une hybridation dans des conditions de forte stringence signifie que les conditions de température et de force ionique sont choisies de telle manière qu'elles permettent le maintien de l'hybridation entre deux fragments d'ADN complémentaires.

- A titre illustratif, des conditions de forte stringence de l'étape d'hybridation aux fins de définir les fragments polynucléotidiques décrits ci-dessus, sont avantageusement les suivantes :

l'hybridation est réalisée à une température préférentielle de 65°C, en présence de tampon commercialisé sous le nom de rapid-hyb buffer par Amersham (RPN 1636) et 100 µg/ml d'ADN de E.coli.

5 Les étapes de lavage peuvent, par exemple, être les suivantes :

- deux lavages de 10 min, préférentiellement à 65°C, dans un tampon 2 x SSC et 0,1% SDS;

- deux lavages de 10 min, préférentiellement à 65°C, dans un 10 tampon 2 x SSC et 0,1% SDS;

- un lavage de 10 min, préférentiellement à 65°C, dans un tampon de 0,1 x SSC et 0,1% SDS.

1 x SSC correspond à 0,15 M NaCl et 0,05M citrate de Na et une solution de 1 x Denhardt correspond à 0,02% Ficoll, 0,02% 15 de polyvinylpyrrolidone et 0,02% de sérum albumine bovine.

Avantageusement, un fragment nucléotidique répondant à la définition précédente aura au moins 8 nucléotides, de préférence au moins 12 nucléotides, et encore plus préférentiellement au moins 20 nucléotides consécutifs de 20 la séquence dont il est issu.

Pour les conditions de mise en oeuvre des enzymes de restriction dans le but d'obtenir des fragments nucléotidiques des polynucléotides selon l'invention, on se référera avantageusement à l'ouvrage de Sambrook de 1989.

25 Avantageusement, un polynucléotide de l'invention contiendra au moins une séquence comprenant l'enchaînement de nucléotides allant du nucléotide en position nt 964 au nucléotide nt 1234 du polynucléotide de séquence SEQ ID N°1.

30 La présente invention a également pour objet un polypeptide issu d'une mycobactérie, caractérisé en ce qu'il est présent uniquement chez les mycobactéries appartenant au complexe de *Mycobacterium tuberculosis*.

La présente invention a aussi pour objet un polypeptide de 35 séquence d'acides aminés SEQ ID N°1 ou SEQ ID N°2.

L'invention concerne également un polypeptide caractérisé en ce qu'il comprend un polypeptide choisi parmi :

- a) un polypeptide de séquence d'acides aminés SEQ ID N°1 ou SEQ ID N°2,
- b) un polypeptide homologue au polypeptide défini en a),
- c) un fragment d'au moins 5 acides aminés d'un polypeptide
- 5 défini en a) ou b),
- d) un fragment biologiquement actif d'un polypeptide défini en a), b), ou c).

Par polypeptide homologue, on entendra désigner les polypeptides présentant, par rapport au polypeptide naturel

10 DP428, certaines modifications comme en particulier une délétion, addition ou substitution d'au moins un acide aminé, une troncation, un allongement, une fusion chimérique, et/ou une mutation. Parmi les polypeptides homologues, on préfère ceux dont la séquence d'acides aminés présente au moins 80%, de

15 préférence 90%, d'homologie avec les séquences d'acides aminés des polypeptides selon l'invention. Dans le cas d'une substitution, un ou plusieurs acides aminés consécutifs ou non consécutifs, sont remplacés par des acides aminés « équivalents ». L'expression acide aminé « équivalent » vise

20 ici à désigner tout acide aminé susceptible d'être substitué à l'un des acides aminés de la structure de base sans cependant modifier essentiellement les propriétés immunogènes des peptides correspondants. En d'autres termes, les acides aminés équivalents seront ceux qui permettent l'obtention d'un

25 polypeptide de séquence modifiée qui permet l'induction *in vivo* d'anticorps capables de reconnaître le polypeptide de séquence SEQ ID N°2 ou l'un de ses fragments ci-dessus définis.

Ces aminoacyles équivalents peuvent être déterminés soit en s'appuyant sur leur homologie de structure avec les

30 aminoacyles auxquels ils se substituent, soit sur les résultats des essais d'immunogénicité croisée auxquels les différents peptides sont susceptibles de donner lieu.

A titre d'exemple, on mentionnera les possibilités de substitutions susceptible d'être effectuées sans qu'il en

35 résulte une modification approfondie de l'immunogénicité des peptides modifiés correspondants, les remplacements, par exemple, de la leucine par la valine ou l'isoleucine, de

l'acide aspartique par l'acide glutamique, de la glutamine par l'asparagine, de l'arginine par la lysine etc., les substitutions inverses étant naturellement envisageables dans les mêmes conditions.

5 Par fragment biologiquement actif, on entendra désigner en particulier un fragment de séquence d'acides aminés de polypeptide présentant au moins une des caractéristiques des polypeptides selon l'invention, notamment en ce qu'il est :

- capable d'être exporté et/ou sécrété par une mycobactérie,
- 10 et/ou d'être induit ou réprimé lors de l'infection par la mycobactérie ; et/ou
- capable d'induire, de réprimer ou de moduler, directement ou indirectement, un facteur de virulence de mycobactérie ; et/ou
- capable d'induire une réaction d'immunogénicité dirigée
- 15 contre les mycobactéries ; et/ou
- capable d'être reconnu par un anticorps spécifique de mycobactérie .

Par fragment de polypeptide, on entend désigner un polypeptide comportant au minimum 5 acides aminés, de

20 préférence 10 acides aminés et 15 acides aminés.

Un polypeptide de l'invention, ou un de ses fragments , tels que définis précédemment, est susceptible d'être reconnu spécifiquement par les anticorps présents dans le sérum de patients infectés par des mycobactéries appartenant au complexe

25 de *Mycobacterium tuberculosis*.

Font ainsi partie de l'invention les fragments du polypeptide DP428 de séquence ID N°1 ou ID N°2, qui peuvent être obtenus par clivage dudit polypeptide par une enzyme protéolytique, telle que la trypsine ou la chymotrypsine ou la

30 collagénase, ou par un réactif chimique, tel que le bromure de cyanogène (CnBr) ou encore en plaçant le polypeptide DP428 dans un environnement très acide, par exemple à pH 2,5.

Des fragments peptidiques préférés selon l'invention, pour une utilisation en diagnostic ou en vaccination, sont les

35 fragments contenus dans des régions du polypeptide DP428 susceptibles d'être naturellement exposées au solvant et de présenter ainsi des propriétés d'immunogénicité importante. De



tels fragments peptidiques peuvent être préparés indifféremment par synthèse chimique, à partir d'hôtes transformés par un vecteur d'expression selon l'invention contenant un acide nucléique permettant l'expression desdits fragments, placé sous le contrôle des éléments de régulation et/ou d'expression appropriés ou encore par clivage chimique ou enzymatique.

Une analyse de l'hydrophilicité du polypeptide DP428 a été réalisée à l'aide du logiciel DNA Strider™ (commercialisé par le CEA Saclay), sur la base d'un calcul du caractère hydrophile de la région codante pour le DP428 de la SEQ ID N°2. Les résultats de cette analyse sont présentés à la figure 4, où sont détaillés, pour chacun des acides aminés (AA) de position 55, 72, 99 et 107, définie dans la SEQ ID N°2, l'indice d'hydrophilicité. Plus l'indice d'hydrophilicité est élevé, plus l'acide aminé considéré est susceptible d'être exposé au solvant dans la molécule native, et est en conséquence susceptible de présenter un degré d'antigénicité élevé. Ainsi, un enchaînement d'au moins sept acides aminés possédant un indice élevé d'hydrophilicité ( $>0,3$ ) peut constituer la base de la structure d'un peptide candidat immunogène selon la présente invention.

L'invention concerne également les séquences d'acide nucléique utilisable comme sonde ou amorce, caractérisées en ce que lesdites séquences sont choisies parmi les séquences d'acide nucléique de polynucléotides selon l'invention.

L'invention concerne en outre l'utilisation d'une séquence d'acide nucléique de polynucléotides selon l'invention comme sonde ou amorce, pour la détection et/ou l'amplification de séquence d'acide nucléique.

Les polynucléotides selon l'invention peuvent ainsi être utilisés pour sélectionner des amorces nucléotidiques, notamment pour la technique PCR.

Cette technique nécessite le choix de paires d'oligonucléotides encadrant le fragment qui doit être amplifié. On peut, par exemple, se référer à la technique décrite dans le brevet américain U.S. N° 4 683 202. Ces amorces oligodésoxyribonucléotidiques ou oligoribonucléotidiques ont

avantageusement une longueur d'au moins 8 nucléotides, de préférence d'au moins 12 nucléotides, et encore plus préférentiellement au moins 20 nucléotides. On préférera en particulier des amorces d'une longueur comprise entre 8 et 30 et de préférence 12 et 22 nucléotides. L'une des deux amorces est complémentaires du brin (+) [amorce aller] de la matrice et l'autre amorce est complémentaire du brin (-) [amorce retour]. Il est important que les amorces ne possèdent pas de structure secondaire ou de séquence complémentaire l'une de l'autre. D'autre part, la longueur et la séquence de chaque amorce doivent être choisies de manière à ce que les amorces ne s'hybrident pas avec d'autres acides nucléiques provenant de cellules procaryotes ou eucaryotes, en particulier avec les acides nucléiques provenant de mycobactéries n'appartenant pas au complexe *Mycobacterium tuberculosis*, ni avec l'ADN ou l'ARN humain pouvant éventuellement contaminer l'échantillon biologique.

Les fragments amplifiés peuvent être identifiés après une électrophorèse en gel d'agarose ou de polyacrylamide, ou après une électrophorèse capillaire, ou encore après une technique chromatographique (filtration sur gel, chromatographie hydrophobe ou chromatographie échangeuse d'ions). La spécificité de l'amplification peut être contrôlée par hybridation moléculaire en utilisant comme sondes les séquences nucléotidiques de polynucléotides de l'invention, des plasmides contenant ces séquences ou leurs produits d'amplification.

Les fragments nucléotidiques amplifiés peuvent être utilisés comme réactifs dans des réactions d'hybridation afin de mettre en évidence la présence, dans un échantillon biologique, d'un acide nucléique cible de séquence complémentaire à celle desdits fragments nucléotidiques amplifiés.

Parmi les polynucléotides selon l'invention, utilisables comme amorces nucléotidiques, on préfère : SEQ ID N°25 et SEQ ID N°26.

Parmi les polynucléotides selon l'invention, utilisables comme sondes nucléotidiques, on préfère le fragment polynucléotidique comprenant les nucléotides 964 à 1234 de SEQ ID N°1.

5 Ces sondes et amplicons peuvent être marqués ou non par des éléments radioactifs ou par des molécules non radioactives, telles que des enzymes ou des éléments fluorescents..

L'invention vise également les fragments nucléotidiques susceptibles d'être obtenus par amplification à l'aide  
10 d'amorces selon l'invention.

D'autres techniques d'amplification de l'acide nucléique cible peuvent être avantageusement employées comme alternatives à la PCR.

La technique SDA (Strand Displacement Amplification) ou  
15 technique d'amplification à déplacement de brin (Walker et al., 1992) est une technique d'amplification isotherme dont le principe est fondé sur la capacité d'une enzyme de restriction de couper l'un des deux brins de son site de reconnaissance qui se trouve sous une forme hemiphosphorothioate et sur la  
20 propriété d'une ADN polymérase d'initier la synthèse d'un nouveau brin d'ADN à partir de l'extrémité 3'OH créée par l'enzyme de restriction et de déplacer le brin préalablement synthétisé qui se trouve en aval.

Les polynucléotides de l'invention, en particulier les  
25 amorces selon l'invention, peuvent également être mis en oeuvre dans d'autres procédés d'amplification d'un acide nucléique cible, tels que :

- la technique TAS (Transcription-based Amplification System), décrite par Kwoh et al. en 1989;
- 30 - la technique 3SR (Self-Sustained Sequence Replication), décrite par Guatelli et al. en 1990;
- la technique NASBA (Nucleic Acid Sequence Based Amplification), décrite par Kievitis et al. en 1991;
- la technique TMA (Transcription Mediated Amplification).

35 Les polynucléotides de l'invention peuvent aussi être employés dans des techniques d'amplification ou de modification de l'acide nucléique servant de sonde, telles que:

- la technique LCR (Ligase Chain Reaction), décrite par Landegren et al. en 1988 et perfectionnée par Barany et al. en 1991, qui emploie une ligase thermostable;
- la technique de RCR (Repair Chain Reaction), décrite par Segev en 1992;
- la technique CPR (Cycling Probe Reaction), décrite par Duck et al. en 1990;
- la technique d'amplification à la Q-beta-réplique, décrite par Miele et al. en 1983 et perfectionnée notamment par Chu et al. en 1986, Lizardi et al. en 1988, puis par Burg et al. ainsi que par Stone et al. en 1996.

Dans le cas où le polynucléotide cible à détecter est un ARN, par exemple un ARNm, on utilisera avantageusement, préalablement à la mise en oeuvre d'une réaction d'amplification à l'aide des amorces selon l'invention ou à la mise en oeuvre d'un procédé de détection à l'aide des sondes de l'invention, une enzyme de type transcriptase inverse afin d'obtenir un ADNc à partir de l'ARN contenu dans l'échantillon biologique. L'ADNc obtenu servira alors de cible pour les amorces ou les sondes mises en oeuvre dans le procédé d'amplification ou de détection selon l'invention.

La sonde de détection sera choisie de telle manière à ce qu'elle hybride avec l'amplicon généré. Une telle sonde de détection aura avantageusement pour séquence une séquence d'au moins 12 nucléotides, en particulier d'au moins 15 nucléotides, et de préférence moins de 200 nucléotides.

Des sondes nucléotidiques selon l'invention sont spécifiques pour détecter les mycobactéries appartenant au complexe de *Mycobacterium tuberculosis*, plus précisément du fait que ces mycobactéries possèdent dans leur génome au moins une copie de polynucléotides selon l'invention. Par les sondes spécifiques selon l'invention, on entend particulièrement tout oligonucléotide hybridant avec la séquence nucléotidique d'un polypeptide selon l'invention, plus particulièrement tout oligonucléotide hybridant avec la séquence SEQ ID N°1 codant pour le polypeptide DP428 de *M. tuberculosis*, et ne présentant pas de réaction d'hybridation croisée ou d'amplification (PCR)

avec des séquences présentes chez des mycobactéries n'appartenant pas au complexe de *Mycobacterium tuberculosis*.

Les sondes nucléotidiques selon l'invention hybrident spécifiquement avec une molécule d'ADN ou d'ARN de polynucléotide selon l'invention, dans des conditions d'hybridation de forte stringence telles que données sous forme d'exemple précédemment.

Les séquences non marquées peuvent être utilisées directement comme sondes, cependant les séquences sont généralement marquées par un élément radioactif ( $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^3\text{H}$ ,  $^{125}\text{I}$ ) ou par une molécule non-radioactive (biotine, acétylaminofluorène, digoxigénine, 5-bromo-désoxyuridine, fluorescéine) pour obtenir des sondes utilisables pour de nombreuses applications.

Des exemples de marquages non radioactifs de sondes sont décrits, par exemple, dans le brevet français N° 78.10975 ou par Urdea et al. ou par Sanchez-Pescador et al. en 1988.

Dans ce dernier cas, on pourra aussi utiliser l'une des méthodes de marquage décrites dans les brevets FR 2 422 956 et FR 2 518 755. La technique d'hybridation peut être réalisée de manières diverses (Matthews et al., 1988). La méthode la plus générale consiste à immobiliser l'acide nucléique extrait des cellules de mycobactéries sur un support (tel que nitrocellulose, nylon, polystyrène) et à incubé, dans des conditions bien définies, l'acide nucléique cible immobilisé avec la sonde. Après l'hybridation, l'excès de sonde est éliminé et les molécules hybrides formées sont détectées par la méthode appropriée (mesure de la radioactivité, de la fluorescence ou de l'activité enzymatique liée à la sonde).

Avantageusement, les sondes nucléotidiques marquées selon l'invention peuvent avoir une structure telle qu'elles rendent possible une amplification du signal radioactif ou non-radioactif. Un système d'amplification répondant à la définition ci-dessus comprendra des sondes de détection sous la forme d'un ADN ramifié, branché (« branched DNA ») telles que celles décrites par Urdea et al. en 1991. Selon cette technique, on utilisera avantageusement plusieurs types de

sondes notamment une sonde de capture, afin d'immobiliser l'ADN ou l'ARN cible sur un support, et une sonde de détection. La sonde de détection lie un ADN « branché » présentant une structure ramifiée. L'ADN branché, à son tour, est capable de  
5 fixer des sondes oligonucléotidiques qui sont elles-mêmes couplées à des molécules de phosphatase alcaline. Puis l'activité de cette enzyme est mise en évidence grâce à un substrat chimioluminescent, par exemple un dérivé du dioxétane-phosphate.

10 Selon un autre mode avantageux de mise en oeuvre des sondes nucléiques selon l'invention, ces dernières peuvent être utilisées comme sondes de capture. Dans ce cas, une sonde, dite « sonde de capture », est immobilisée sur un support de manière covalente ou non covalente et sert à capturer par hybridation  
15 spécifique l'acide nucléique cible obtenu à partir de l'échantillon biologique à tester. Si nécessaire, le support solide est séparé de l'échantillon et le duplex formé entre la sonde de capture et l'acide nucléique cible est ensuite détecté grâce à une seconde sonde, dite « sonde de détection », marquée  
20 par un élément facilement détectable.

Les fragments oligonucléotidiques peuvent être obtenus à partir des séquences selon l'invention, par coupure avec des enzymes de restriction, ou par synthèse chimique selon les méthodes classiques, par exemple selon la méthode décrite dans  
25 le brevet européen N° EP-0305929 (Millipore Corporation) ou encore par d'autres procédés.

Un mode de préparation approprié des acides nucléiques de l'invention comportant au maximum 200 nucléotides (ou 200 pb s'il s'agit d'acides nucléiques bicaténaires) comprend les  
30 étapes suivantes :

- la synthèse d'ADN en utilisant la méthode automatisée des bêta-cyanethylphosphoramidite décrite en 1986,
- le clonage des acides nucléiques ainsi obtenus dans un vecteur approprié et la récupération de l'acide nucléique par  
35 hybridation avec une sonde appropriée.

Un mode de préparation, par voie chimique, d'acides nucléiques selon l'invention de longueur supérieure à 200

nucléotides (ou 200 pb lorsqu'il s'agit d'acides nucléiques bicaténaires) comprend les étapes suivantes :

- l'assemblage d'oligonucléotides synthétisés chimiquement, pourvus à leur extrémité de sites de restrictions différents, dont les séquences sont compatibles avec l'enchaînement en acides aminés du polypeptide naturel selon le principe décrit en 1983,
- le clonage des acides nucléiques ainsi obtenus dans un vecteur approprié et la récupération de l'acide nucléique recherché par hybridation avec une sonde appropriée.

Les sondes nucléotidiques utilisées pour la récupération de l'acide nucléique recherché dans les procédés susmentionnés, sont constituées généralement de 8 à 200 nucléotides de la séquence de polypeptide selon l'invention et sont susceptibles de s'hybrider avec l'acide nucléique recherché dans les conditions d'hybridation définies précédemment. La synthèse de ces sondes peut être effectuée selon la méthode automatisée des bêta cyanethylphosphoramidites décrite en 1986.

Les sondes oligonucléotidiques selon l'invention peuvent être mises en oeuvre au sein d'un dispositif de détection comprenant une banque matricielle d'oligonucléotides. Un exemple de réalisation d'une telle banque matricielle peut consister en une matrice d'oligonucléotides sondes fixés sur un support, la séquence de chaque sonde d'une longueur donnée étant située en décalage d'une ou plusieurs bases par rapport à la sonde précédente, chacune des sondes de l'arrangement matriciel étant ainsi complémentaire d'une séquence distincte de l'ADN ou l'ARN cible à détecter et chaque sonde de séquence connue étant fixée en une position prédéterminée du support. La séquence cible à détecter peut être avantageusement marquée radioactivement ou non radioactivement. Lorsque la séquence cible marquée est mise en contact avec le dispositif matriciel, celle-ci forme des hybrides avec les sondes de séquences complémentaires. Un traitement à la nucléase, suivi d'un lavage, permet d'éliminer les hybrides sondes-séquence cible qui ne sont pas parfaitement complémentaires. Du fait de la

connaissance précise de la séquence d'une sonde à une position déterminée de la matrice, il est alors possible de déduire la séquence nucléotidique de la séquence d'ADN ou d'ARN cible. Cette technique est particulièrement efficace lorsque sont  
5 utilisées des matrices de sondes oligonucléotidiques de grande taille.

Une alternative à l'utilisation d'une séquence cible marquée peut consister en l'utilisation d'un support permettant une détection « bioélectronique » de l'hybridation de la  
10 séquence cible sur les sondes du support matrice, lorsque que ledit support est constitué ou comprend un matériau capable d'agir, par exemple, en tant que donneur d'électrons aux positions de la matrice auxquelles un hybride a été formé. Un tel matériau donneur d'électron est par exemple de l'or. La  
15 détection de la séquence nucléotidique de l'ADN ou ARN cible est alors déterminée par un dispositif électronique.

Un exemple de réalisation d'un biocapteur, tel que défini ci-dessus, est décrit dans la demande de brevet européen N° EP-0721 016 au nom de Affymax technologies N.V. ou encore dans le  
20 brevet américain N° US 5.202.231 au nom de Drmanac.

L'invention a aussi pour objet les polynucléotides hybrides résultant :

- soit de la formation d'une molécule hybride entre un ARN ou un ADN (ADN génomique ou ADNc) provenant d'un échantillon  
25 biologique avec une sonde ou une amorce selon l'invention.
- soit de la formation d'une molécule hybride entre un ARN ou un ADN (ADN génomique ou ADNc) provenant d'un échantillon biologique avec un fragment nucléotidique amplifié à l'aide d'un couple d'amorces selon l'invention.

30 Par ADNc au sens de la présente invention, on entend une molécule d'ADN obtenue en faisant agir une enzyme de type transcriptase inverse sur une molécule d'ARN, en particulier une molécule d'ARN messenger (ARNm), selon les techniques décrites dans Sambrook et al. en 1989.

35 La présente invention a également pour objet une famille de plasmides recombinés, caractérisés en ce qu'ils contiennent au moins une séquence nucléotidique de polynucléotide selon



l'invention. Selon un mode de réalisation avantageux dudit plasmide, il comprend la séquence nucléotidique SEQ ID N°1 ou un fragment de celle-ci.

Un autre objet de la présente invention est un vecteur  
5 pour le clonage, l'expression et/ou l'insertion d'une séquence, caractérisé en ce qu'il comprend une séquence nucléotidique de polynucléotide selon l'invention en un site non essentiel pour sa répllication, le cas échéant sous le contrôle d'éléments de régulation susceptibles d'intervenir dans l'expression du  
10 polypeptide DP428, chez un hôte déterminé.

Des vecteurs particuliers sont par exemple des plasmides, des phages, des cosmides, des phagemides, des YAC.

Ces vecteurs sont utiles pour transformer des cellules hôtes afin de cloner ou d'exprimer les séquences nucléotidiques  
15 de l'invention.

L'invention comprend également les cellules hôtes transformées par un vecteur selon l'invention.

De préférence, les cellules hôtes sont transformées dans des conditions permettant l'expression d'un polypeptide  
20 recombinant selon l'invention.

Une cellules hôte préférée selon l'invention est une mycobactérie appartenant à une souche de *M. tuberculosis*, *M.bovis* ou *M.africanum* possédant potentiellement tous les systèmes de régulation appropriés.

25 Il est aujourd'hui facile de produire des protéines ou polypeptides en quantité relativement importante par génie génétique en utilisant comme vecteurs d'expression des plasmides, des phages, des phagemides. Tout ou partie du gène DP428, ou tout polynucléotide selon l'invention, peut être  
30 inséré dans un vecteur d'expression approprié pour produire *in vitro* un polypeptide selon l'invention, notamment le polypeptide DP428. Ledit polypeptide pourra être fixé sur une microplaque pour développer un test sérologique destiné à rechercher, dans un but de diagnostic, les anticorps  
35 spécifiques chez les patients atteints de tuberculose.

Ainsi, la présente invention concerne plus particulièrement un procédé de préparation d'un polypeptide de l'invention comprenant les étapes suivantes :

- 5 - le cas échéant, l'amplification préalable suivant la technique PCR de la quantité de séquences de nucléotides codant pour ledit polypeptide à l'aide de deux amorces d'ADN choisies de manière à ce que l'une de ces amorces soit identique aux 10 à 25 premiers nucléotides de la séquence nucléotidique codant pour ledit polypeptide, tandis que l'autre amorce est  
10 complémentaire des 10 à 25 derniers nucléotides (ou s'hybride avec ces 10 à 25 derniers nucléotides) de ladite séquence nucléotidique, ou inversement de manière à ce que l'une de ces amorces soit identique aux 10 à 25 derniers nucléotides de ladite séquence, tandis que l'autre amorce est complémentaire  
15 des 10 à 25 premiers nucléotides (ou s'hybride avec les 10 à 25 premiers nucléotides) de ladite séquence nucléotidique, suivie de l'introduction desdites séquences ainsi amplifiées dans un vecteur approprié,
- la mise en culture, dans un milieu de culture approprié, d'un  
20 hôte cellulaire préalablement transformé par un vecteur approprié contenant un acide nucléique selon l'invention comprenant la séquence nucléotidique codant pour ledit polypeptide, et
- la séparation, à partir du susdit milieu de culture, dudit  
25 polypeptide produit par ledit hôte cellulaire transformé.

L'invention a aussi pour objet un polypeptide, caractérisé en ce qu'il est susceptible d'être obtenu par un procédé de l'invention tel que décrit précédemment.

- 30 Les peptides selon l'invention peuvent également être préparés par les techniques classiques, dans le domaine de la synthèse des peptides. Cette synthèse peut être réalisée en solution homogène ou en phase solide.

Par exemple, on aura recours à la technique de synthèse en  
35 solution homogène décrite par Houbenweyl en 1974.

Cette méthode de synthèse consiste à condenser successivement deux-à-deux les aminoacyles successifs dans

l'ordre requis, ou à condenser des aminoacyles et des fragments préalablement formés et contenant déjà plusieurs aminoacyles dans l'ordre approprié, ou encore plusieurs fragments préalablement ainsi préparés, étant entendu que l'on aura eu  
5 soin de protéger au préalable toutes les fonctions réactives portées par ces aminoacyles ou fragments, à l'exception des fonctions amines de l'un et carboxyles de l'autre ou vice-versa, qui doivent normalement intervenir dans la formation des liaisons peptidiques, notamment après activation de la fonction  
10 carboxyle, selon les méthodes bien connues dans la synthèse des peptides. En variante, on pourra avoir recours à des réactions de couplage mettant en jeu des réactifs de couplage classique, du type carbodiimide, tels que par exemple la 1-éthyl-3-(3-diméthyl-aminopropyl)-carbodiimide.

15 Lorsque l'aminoacyle mis en oeuvre possède une fonction acide supplémentaire (notamment dans le cas de l'acide glutamique), ces fonctions seront protégées, par exemple par des groupes t-butylester.

Dans le cas de la synthèse progressive, acide aminé par  
20 acide aminé, la synthèse débute de préférence par la condensation de l'amino-acide C-terminal avec l'aminoacide qui correspond à l'aminoacyle voisin dans la séquence désirée et ainsi de suite, de proche en proche, jusqu'à l'acide aminé N-terminal.

25 Selon une autre technique préférée de l'invention, on a recours à celle décrite par Merrifield.

Pour fabriquer une chaîne peptidique selon le procédé de Merrifield, on a recours à une résine polymère très poreuse, sur laquelle on fixe le premier acide aminé C-terminal de la  
30 chaîne. Cet acide aminé est fixé sur la résine par l'intermédiaire de son groupe carboxylique et sa fonction amine est protégée, par exemple par le groupe t-butyloxycarbonyle.

Lorsque le premier acide aminé C-terminal est ainsi fixé sur la résine, on enlève le groupe protecteur de la fonction  
35 amine en lavant la résine avec un acide.

Dans le cas où le groupe protecteur de la fonction amine est le groupe t-butyloxycarbonyle, il peut être éliminé par traitement de la résine à l'aide d'acide trifluoroacétique.

On couple ensuite le deuxième acide aminé qui fournit le  
5 second aminoacyle de la séquence recherchée, à partir du résidu aminoacyle C-terminal sur la fonction amine déprotégée du premier acide aminé C-terminal fixé sur la chaîne. De préférence, la fonction carboxyle de ce deuxième acide aminé est activée, par exemple par la dicyclohexylcarbodiimide, et la  
10 fonction amine est protégée, par exemple par le t-butyloxycarbonyle.

On obtient ainsi la première partie de la chaîne peptidique recherchée, qui comporte deux acides aminés, et dont la fonction amine terminale est protégée. Comme précédemment,  
15 on déprotège la fonction amine et on peut ensuite procéder à la fixation du troisième aminoacyle, dans des conditions analogues à celles de l'addition du deuxième acide aminé C-terminal.

On fixe ainsi, les uns après les autres, les acides aminés qui vont constituer la chaîne peptidique sur le groupe amine  
20 chaque fois déprotégé au préalable de la portion de la chaîne peptidique déjà formée, et qui est rattachée à la résine.

Lorsque la totalité de la chaîne peptidique désirée est formée, on élimine les groupes protecteurs des différents acides aminés constituant la chaîne peptidique et on détache le  
25 peptide de la résine, par exemple à l'aide d'acide fluorhydrique.

De manière préférentielle, lesdits polypeptides susceptibles d'être obtenus par un procédé de l'invention tel que décrit précédemment comprendront une région exposée au  
30 solvant et auront une longueur d'au moins 20 acides aminés.

Selon un autre mode de réalisation de l'invention, lesdits polypeptides sont spécifiques de mycobactéries du complexe *Mycobacterium tuberculosis* et ne sont donc pas reconnus par des anticorps spécifiques d'autres protéines de mycobactéries.

35 L'invention est en outre relative à des polypeptides hybrides présentant au moins un polypeptide selon l'invention

et une séquence d'un polypeptide susceptible d'induire une réponse immunitaire chez l'homme ou l'animal.

Avantageusement, le déterminant antigénique est tel qu'il est susceptible d'induire une réponse humorale et/ou cellulaire, comme par  
5 exemple les déterminants antigéniques de protéines immunogènes sélectionnées par exemple parmi ESAT 6 et DES de 45/47 kD de *Mycobacterium tuberculosis*.

Un tel déterminant pourra comprendre un polypeptide selon l'invention sous forme glycosylée utilisé en vue d'obtenir des compositions immunogènes  
10 susceptibles d'induire la synthèse d'anticorps dirigés contre des épitopes multiples. Lesdits polypeptides glycosylés font également partie de l'invention.

Ces molécules hybrides peuvent être constituées en partie d'une molécule porteuse de polypeptide selon l'invention associée à une partie,  
15 en particulier un épitope de la toxine diphtérique, la toxine tétanique, un antigène de surface du virus de l'hépatite B (brevet FR 79 21811), l'antigène VP1 du virus de la poliomyélite ou toute autre toxine ou antigène viral ou bactérien.

Avantageusement, un antigène bactérien tel que défini ci-dessus sera  
20 tout ou partie de la protéine immunogène de 45/47 kD de *M. tuberculosis* (demande internationale PCT/FR 96/0166).

Un antigène viral, tel que défini ci-dessus, sera préférentiellement une protéine de surface ou d'enveloppe d'un virus de l'hépatite, par exemple la protéine de surface de l'hépatite B sous l'une de ses formes S,  
25 S-prés1, S-prés2 ou S-prés2-prés1 ou encore une protéine d'un virus de l'hépatite A, ou d'une hépatite non-A, non-B, tel qu'un virus de l'hépatite C, E ou delta.

Plus particulièrement, un antigène viral tel que défini ci-dessus sera tout ou partie de l'une des glycoprotéines codées par le génome du  
30 virus HIV-1 (brevets GB 8324800, EP 84401834 ou EP 85905513) ou du virus HIV-2 (EP 87400151), et en particulier tout ou partie d'une protéine sélectionnée parmi gag, pol, nef ou env de HIV-1 ou de HIV-2.

Les procédés de synthèse des molécules hybrides englobent les méthodes utilisées en génie génétique pour construire des ADN hybrides  
35 codant pour les séquences polypeptidiques recherchées. On pourra, par exemple, se référer avantageusement à la technique d'obtention de gènes codant pour des protéines de fusion décrite par Minton en 1984.

Les polynucléotides hybrides codant pour un polypeptide hybride ainsi que les polypeptides hybrides selon l'invention caractérisés en ce qu'il s'agit de protéines recombinantes obtenues par l'expression desdits polynucléotides hybrides, font également partie de l'invention.

5 Les polypeptides selon l'invention peuvent avantageusement être mis en oeuvre dans un procédé de détection *in vitro* d'anticorps dirigés contre lesdits polypeptides, notamment le polypeptide DP428, et ainsi contre une bactérie du complexe *M. tuberculosis* dans un échantillon biologique (tissu ou fluide biologique) susceptible de les contenir, ce procédé comprenant la  
10 mise en contact de cet échantillon biologique avec un polypeptide selon l'invention dans des conditions permettant une réaction immunologique *in vitro* entre ledit polypeptide et les anticorps éventuellement présents dans l'échantillon biologique, et la détection *in vitro* des complexes antigène-anticorps éventuellement formés.

15 De préférence, l'échantillon biologique est constitué par un fluide, par exemple un sérum humain ou animal.

Toute procédure classique peut être mise en oeuvre pour réaliser une telle détection.

A titre d'exemple, une méthode préférée met en jeu des  
20 processus immunoenzymatiques selon la technique ELISA, par immunofluorescence, ou radio-immunologique (RIA) ou équivalent.

Ainsi, l'invention concerne également les polypeptides selon l'invention, marqués à l'aide d'un marqueur adéquat tel que du type enzymatique, fluorescent, radioactif.

25 De telles méthodes comprennent par exemple les étapes suivantes :

- dépôt de quantités déterminées d'une composition polypeptidique selon l'invention dans les puits d'une plaque de microtitration,
- 30 - introduction dans lesdits puits de dilutions croissantes du sérum devant être analysé,
- incubation de la microplaque,
- introduction dans les puits de la plaque de microtitration d'anticorps marqués dirigés contre des immunoglobulines  
35 humaines ou animales, le marquage de ces anticorps ayant été réalisé à l'aide d'une enzyme sélectionnée parmi celles qui sont capables d'hydrolyser un substrat en modifiant

l'absorption des radiations de ce dernier, au moins à une longueur d'onde déterminée, par exemple à 550 nm, -détection, en comparaison avec un témoin de contrôle, de la quantité de substrat hydrolysé.

5 L'invention concerne également un nécessaire ou kit pour le diagnostic in vitro d'une infection par une mycobactérie appartenant au complexe *Mycobacterium tuberculosis*, comprenant:

- un polypeptide selon l'invention,
- les réactifs pour la constitution du milieu propice à la
- 10 réaction immunologique,
- les réactifs permettant la détection des complexes antigène-anticorps produits par la réaction immunologique, ces réactifs peuvent également porter un marqueur, ou être susceptibles d'être reconnus à leur tour par un réactif marqué, plus
- 15 particulièrement dans le cas où le polypeptide selon l'invention n'est pas marqué,
- le cas échéant, un échantillon biologique de référence (témoin négatif) dépourvu d'anticorps reconnus par un polypeptide selon l'invention,
- 20 - le cas échéant, un échantillon biologique de référence (témoin positif) contenant une quantité prédéterminée d'anticorps reconnus par un polypeptide selon l'invention.

Les polypeptides selon l'invention permettent de préparer des anticorps monoclonaux ou polyclonaux caractérisés en ce

25 qu'ils reconnaissent spécifiquement les polypeptides selon l'invention. Les anticorps monoclonaux pourront avantageusement être préparés à partir d'hybridomes selon la technique décrite par Kohler et Milstein en 1975. Les anticorps polyclonaux pourront être préparés, par exemple par immunisation d'un

30 animal, en particulier une souris, avec un polypeptide selon l'invention associé à un adjuvant de la réponse immunitaire, puis purification des anticorps spécifiques contenus dans le sérum des animaux immunisés sur une colonne d'affinité sur laquelle a préalablement été fixé le polypeptide ayant servi

35 d'antigène. Les anticorps polyclonaux selon l'invention peuvent aussi être préparés par purification sur une colonne d'affinité, sur laquelle a préalablement été immobilisé un

polypeptide selon l'invention, des anticorps contenus dans le sérum de patients infectés par une mycobactérie appartenant au complexe *Mycobacterium tuberculosis*.

5 L'invention a également pour objet des anticorps mono ou polyclonaux ou leurs fragments, ou anticorps chimériques, caractérisés en ce qu'ils sont capables de reconnaître spécifiquement un polypeptide selon l'invention.

10 Les anticorps de l'invention pourront également être marqués de la même manière que décrit précédemment pour les sondes nucléiques de l'invention tel qu'un marquage de type enzymatique, fluorescent ou radioactif.

L'invention vise en outre un procédé pour la détection spécifique de la présence d'un antigène d'une bactérie du complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

- a) Mise en contact de l'échantillon biologique (tissu ou fluide biologique) prélevé chez un individu avec un anticorps mono ou polyclonal selon l'invention, dans des conditions permettant une réaction immunologique *in vitro* entre lesdits anticorps et les polypeptides spécifiques des mycobactéries du complexe de *Mycobacterium tuberculosis* éventuellement présents dans l'échantillon biologique, et
- 25 b) Mise en évidence du complexe antigène-anticorps formé.

Entre également dans le cadre de l'invention, un nécessaire ou kit pour le diagnostic *in vitro* sur un échantillon biologique, de la présence de souches de mycobactéries appartenant au complexe de *Mycobacterium tuberculosis*, de préférence *M. tuberculosis*, caractérisé en ce qu'il comprend :

- un anticorps polyclonal ou monoclonal selon l'invention, le cas échéant marqué;
- 35 - le cas échéant, un réactif pour la constitution du milieu propice à la réalisation de la réaction immunologique;



- un réactif permettant la détection des complexes antigène-anticorps produits par la réaction immunologique, ce réactif pouvant également porter un marqueur, ou être susceptible d'être reconnu à son tour par un réactif marqué, plus particulièrement dans le cas où ledit anticorps monoclonal ou polyclonal n'est pas marqué.
- le cas échéant, des réactifs pour effectuer la lyse des cellules de l'échantillon testé.

10

La présente invention a également pour objet un procédé de détection et d'identification rapide de *M. tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes :

- 15 a) Isolement de l'ADN à partir de l'échantillon biologique à analyser, ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ;
- b) Amplification spécifique de l'ADN des mycobactéries appartenant au complexe *Mycobacterium tuberculosis* à l'aide d'amorces selon l'invention;
- 20 c) Analyse des produits d'amplification.

Ceux-ci peuvent être analysés par différentes méthodes.

Deux méthodes d'analyse sont données à titre d'exemple ci-dessous :

- 25 - Analyse électrophorétique en gel d'agarose des produits d'amplification. Si l'on observe la présence d'un fragment d'ADN migrant à l'endroit attendu, on peut conclure que l'échantillon analysé contenait de l'ADN de mycobactéries appartenant au complexe tuberculosis, ou
- 30 - Analyse par la technique d'hybridation moléculaire en utilisant une sonde nucléique selon l'invention. Cette sonde sera avantageusement marquée par un élément non radioactif (sonde froide) ou radioactif.

Aux fins de la présente invention, on entendra par « ADN de l'échantillon biologique » ou « ADN contenu dans l'échantillon biologique », soit l'ADN présent dans

l'échantillon biologique considéré, soit l'ADNc obtenu après l'action d'une enzyme de type transcriptase inverse sur l'ARN présent dans ledit échantillon biologique.

Un autre but de la présente invention consiste en un  
5 procédé pour la détection spécifique de bactéries appartenant au complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

a) Mise en contact d'une sonde oligonucléotidique selon  
10 l'invention avec un échantillon biologique, l'ADN contenu dans l'échantillon biologique, ou l'ADNc obtenu par transcription inverse de l'ARN de l'échantillon biologique, ayant, le cas échéant, préalablement été rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de la sonde à  
15 l'ADN ou l'ADNc d'une bactérie du complexe *Mycobacterium tuberculosis*;

b) Détection de l'hybride formé entre la sonde oligonucléotidique et l'ADN de l'échantillon biologique.

20 L'invention vise également un procédé pour la détection spécifique de bactéries appartenant au complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

a) Mise en contact d'une sonde oligonucléotidique selon  
25 l'invention immobilisée sur un support, avec un échantillon biologique, l'ADN de l'échantillon biologique ayant, le cas échéant, été préalablement rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de ladite sonde à l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*;

30 b) Mise en contact de l'hybride formé entre ladite sonde oligonucléotidique immobilisée sur un support et l'ADN contenu dans l'échantillon biologique, le cas échéant après élimination de l'ADN de l'échantillon biologique n'ayant pas hybridé avec la sonde, avec une sonde oligonucléotidique marquée selon  
35 l'invention.

Selon un mode de réalisation avantageux du procédé de détection défini précédemment, celui-ci est caractérisé en ce que, préalablement à l'étape a), l'ADN de l'échantillon biologique est préalablement amplifié à l'aide d'un couple d'amorces selon l'invention.

Une autre forme de mise en oeuvre du procédé de détection selon l'invention consiste en un procédé pour la détection spécifique de la présence d'une bactérie appartenant au complexe de *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

- a) Mise en contact de l'échantillon biologique avec un couple d'amorces selon l'invention, l'ADN contenu dans l'échantillon ayant été, le cas échéant, préalablement rendu accessible à l'hybridation, dans des conditions permettant une hybridation desdites amorces à l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*;
- b) Amplification de l'ADN de la bactérie du complexe *Mycobacterium tuberculosis*;
- c) mise en évidence de l'amplification de fragments d'ADN correspondant au fragment encadré par les amorces, par exemple par électrophorèse sur gel ou au moyen d'une sonde oligonucléotidique selon invention.

L'invention a aussi pour objet un procédé pour la détection spécifique de la présence d'une bactérie appartenant au complexe de *Mycobacterium tuberculosis* dans un échantillon biologique par déplacement de brin, caractérisé en ce qu'il comprend les étapes suivantes :

- a) Mise en contact de l'échantillon biologique avec deux couples d'amorces selon l'invention spécifiquement destinées à l'amplification de type SDA décrites ci-dessus, l'ADN contenu dans l'échantillon ayant été, le cas échéant, préalablement rendu accessible à l'hybridation, dans des conditions permettant une hybridation des amorces à l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*;

b) amplification de l'ADN de la bactérie du complexe *Mycobacterium tuberculosis*;

c) mise en évidence de l'amplification de fragments d'ADN correspondant au fragment encadré par les amorces, par exemple  
5 par électrophorèse sur gel ou au moyen d'une sonde oligonucléotidique selon l'invention.

L'invention concerne aussi un nécessaire ou kit pour la mise en oeuvre du procédé décrit ci-dessus, destiné à la détection de la présence d'une bactérie du complexe  
10 *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants :

- a) Une sonde oligonucléotidique selon l'invention;
- b) Les réactifs nécessaires à la mise en oeuvre d'une réaction d'hybridation;
- 15 c) Le cas échéant, un couple d'amorces selon l'invention ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN (ADN génomique, ADN plasmidique ou ADNc) d'une bactérie du complexe *Mycobacterium tuberculosis*.

L'invention a aussi pour objet un kit ou nécessaire pour  
20 la détection de la présence d'une bactérie du complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants :

- a) Une sonde oligonucléotidique, dite sonde de capture, selon l'invention;
- 25 b) Une sonde oligonucléotidique, dite sonde de révélation, selon l'invention.
- c) Le cas échéant, un couple d'amorces selon l'invention ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*.

30 L'invention concerne encore un kit ou nécessaire pour l'amplification de l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis* présent dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants :

- 35 a) Un couple d'amorces selon l'invention;

b) Les réactifs nécessaires pour effectuer une réaction d'amplification d'ADN;

c) Eventuellement un composant permettant de vérifier la séquence du fragment amplifié, plus particulièrement une sonde  
5 oligonucléotidique selon l'invention.

Un autre objet de la présente invention concerne une composition immunogène, caractérisée en ce qu'elle comprend polypeptide selon l'invention.

Une autre composition immunogène selon l'invention est  
10 caractérisé en ce qu'elle comprend un ou plusieurs polypeptides selon l'invention et/ou une ou plusieurs protéines hybrides selon l'invention.

Selon un mode de réalisation avantageux, la composition immunogène ci-dessus définie est constitutive d'un vaccin,  
15 lorsqu'elle est présentée en association avec un véhicule pharmaceutiquement acceptable et éventuellement un ou plusieurs adjuvants de l'immunité tels que l'alun ou un représentant de la famille des muramyl peptides ou encore l'adjuvant incomplet de Freund.

20 Aujourd'hui, divers types vaccins sont disponibles pour protéger l'homme contre des maladies infectieuses : micro-organismes vivants atténués (*M. bovis* - BCG pour la tuberculose), micro-organismes inactivés (virus de la grippe), des extraits acellulaires (*Bordetella pertussis* pour la  
25 coqueluche), protéines recombinées (antigène de surface du virus de l'hépatite B), des polysides (pneumocoques). Des vaccins préparés à partir de peptides de synthèse ou de micro-organismes génétiquement modifiés exprimant des antigènes hétérologues sont en cours d'expérimentation. Plus récemment  
30 encore, des ADN plasmidiques recombinés portant des gènes codant pour des antigènes protecteurs ont été proposés comme stratégie vaccinale alternative. Ce type de vaccination est réalisé avec un plasmide particulier dérivant d'un plasmide de *E. coli* qui ne se réplique pas *in vivo* et qui code  
35 uniquement pour la protéine vaccinante. Les principaux composants fonctionnels de ce plasmide sont : un promoteur fort (par exemple celui du CMV), un site de clonage approprié

pour insérer le gène d'intérêt, une séquence de terminaison-polyadénylation, une origine de répllication procaryote pour produire le plasmide recombiné *in vitro* et un marqueur de sélection (par exemple le gène de résistance à l'ampicilline) pour faciliter la sélection des bactéries qui contiennent le plasmide. Des animaux ont été immunisés en injectant simplement l'ADN plasmidique nu dans le muscle. Cette technique conduit à l'expression de la protéine vaccinale *in situ* et à une réponse immunitaire de type cellulaire (CTL) et de type humoral (anticorps). Cette double induction de la réponse immunitaire est l'un des principaux avantages de la technique de vaccination avec de l'ADN nu. Huygen et al. (1996) et Tascon et al. (1996) ont réussi à obtenir une certaine protection contre *M. tuberculosis* en injectant des plasmides recombinés contenant des gènes de *M. leprae* (*hsp65*, *36kDa pra*) comme inserts. *M. leprae* est l'agent responsable de la lèpre. L'utilisation d'un insert spécifique de *M. tuberculosis* comme par exemple tout ou partie du gène *DP428*, objet de la présente invention conduirait probablement à une meilleure protection contre la tuberculose. Tout ou partie du gène *DP428*, ou tout polynucléotide selon l'invention, peut être facilement inséré dans les plasmides vecteurs *V1J* (Montgomery et al, 1993), *pcDNA3* (Invitrogen, R & D Systems) ou *pcDNA1/Neo* (Invitrogen) qui possèdent les caractéristiques nécessaires pour une utilisation vaccinale.

L'invention vise ainsi une composition immunogène caractérisée en ce qu'elle comprend un ou plusieurs polypeptides hybrides tels que précédemment définis en association avec un véhicule pharmaceutiquement compatible et, le cas échéant, un ou plusieurs adjuvants de l'immunité.

L'invention vise aussi une composition vaccinale destinée à l'immunisation de l'homme ou l'animal à l'encontre d'une infection bactérienne ou virale, telle que la tuberculose ou l'hépatite, caractérisée en ce qu'elle comprend un ou plusieurs polypeptides hybrides tels que précédemment définis en

association avec un véhicule pharmaceutiquement compatible et, le cas échéant, un ou plusieurs adjuvants de l'immunité.

Avantageusement, dans le cas d'une protéine hybride entre un polypeptide selon l'invention et l'antigène de surface de l'hépatite B, la composition vaccinale sera administrée, chez  
5 l'homme, à raison de 0,1 à 1  $\mu$ g de protéine hybride purifiée par kilogramme du poids du patient, de préférence 0,2 à 0,5  $\mu$ g/kg de poids du patient, pour une dose destinée à une administration donnée. Dans le cas de patients atteints de  
10 troubles du système immunitaire, en particulier les patients immunodéprimés, chaque dose injectée contiendra préférentiellement la moitié de la quantité pondérale de la protéine hybride contenue dans une dose destinée à un patient n'étant pas affecté de troubles du système immunitaire.

15 De préférence, la composition vaccinale sera administrée à plusieurs reprises, de manière étalée dans le temps, par voie intradermique ou sous-cutanée. A titre d'exemple, trois doses telles que définies ci-dessus seront respectivement administrées au patient au temps  $t_0$ , au temps  $t_0 + 1$  mois et au  
20 temps  $t_0 + 1$  an.

Alternativement, trois doses seront respectivement administrées au patient au temps  $t_0$ , au temps  $t_0 + 1$  mois et au temps  $t_0 + 6$  mois.

Chez la souris, chez laquelle une dose pondérale de la  
25 composition vaccinale comparable à la dose utilisée chez l'homme est administrée, la réaction anticorps est testée par prélèvement du sérum suivi d'une étude de la formation d'un complexe entre les anticorps présents dans le sérum et l'antigène de la composition vaccinale, selon les techniques  
30 usuelles.

L'invention concerne également une composition immunogène caractérisée en ce qu'elle comprend un polynucléotide ou un vecteur d'expression selon l'invention, en association avec un véhicule permettant son administration à l'homme ou l'animal.

35 L'invention a encore pour objet un vaccin destiné à l'immunisation à l'encontre d'une infection bactérienne ou virale, caractérisé en ce qu'il comprend un polynucléotide ou

un vecteur d'expression selon l'invention, en association avec un véhicule pharmaceutiquement acceptable.

De telles compositions immunogènes ou vaccinales sont notamment décrites dans la demande internationale N° WO 90/11092 (Vical Inc.) et également dans la demande internationale N° WO 95/11307 (Institut Pasteur).

Le polynucléotide constitutif de la composition immunogène ou de la composition vaccinale selon l'invention peut être injecté à l'hôte après avoir été couplé à des composés qui favorisent la pénétration de ce polynucléotide à l'intérieur de la cellule ou son transport jusqu'au noyau cellulaire. Les conjugués résultants peuvent être encapsulés dans des microparticules polymères, comme décrit dans la demande internationale N° WO 94/27238 (medisorb Technologies International).

Selon un autre mode de réalisation de la composition immunogène et/ou vaccinale selon l'invention, le polynucléotide, de préférence un ADN, est complexé avec du DEAE-dextran (Pagano et al., 1967) ou avec des protéines nucléaires (Kaneda et al., 1989), avec des lipides (Felgner et al., 1987) ou encore encapsulés dans des liposomes (Fraley et al., 1980).

Selon encore un autre mode de réalisation avantageux de la composition immunogène et/ou vaccinale selon l'invention, le polynucléotide selon l'invention peut être introduit sous la forme d'un gel facilitant sa transfection dans les cellules. Une telle composition sous forme de gel peut être un complexe de poly-L-lysine et de lactose, comme décrit par Midoux en 1993, ou encore le Poloxamer 407™, comme décrit par Pastore en 1994. Le polynucléotide ou le vecteur selon l'invention peuvent aussi être en suspension dans une solution tampon ou être associés à des liposomes.

Avantageusement, un tel vaccin sera préparé conformément à la technique décrite par Tacson et al. ou Huygen et al. en 1996 ou encore conformément à la technique décrite par Davis et al. dans la demande internationale N° WO 95/11307 (Whalen et al.).



Un tel vaccin sera avantageusement préparé sous la forme d'une composition contenant un vecteur selon l'invention, placée sous le contrôle d'éléments de régulation permettant son expression chez l'homme ou l'animal.

5 Pour réaliser un tel vaccin, le polynucléotide selon l'invention est tout d'abord sous-cloné dans un vecteur d'expression approprié, plus particulièrement un vecteur d'expression contenant des signaux de régulation et d'expression reconnus par les enzymes des cellules eucaryotes  
10 et contenant également une origine de répllication active chez les procaryotes, par exemple chez *E. coli*, qui permet son amplification préalable. Puis le plasmide recombinant purifié obtenu est injecté à l'hôte, par exemple par voie intramusculaire.

15 On pourra par exemple utiliser, en tant que vecteur d'expression in vivo de l'antigène d'intérêt, le plasmide pcDNA3 ou le plasmide pcDNA1/neo, tous les deux commercialisés par Invitrogen (R&D Systems, Abingdon, Royaume-Uni). On peut aussi utiliser le plasmide VIJns.tPA, décrit par Shiver et al.  
20 en 1995.

Un tel vaccin comprendra avantageusement, outre le vecteur recombinant, une solution saline, par exemple une solution de chlorure de sodium.

25 Une composition vaccinale telle que définie ci-dessus sera par exemple administrée par voie parentérale ou par voie intramusculaire.

D'autres caractéristiques et avantages de l'invention apparaissent dans les exemples et les figures suivants :

30

#### FIGURES

##### Figure 1 :

1A. la construction pJVED: Plasmid navette( pouvant se  
35 multiplier chez les mycobactéries ainsi que chez *E.coli*). avec un gène de résistance à la kanamycine (issu de Tn903)

comme marqueur de sélection. Le gène *phoA* tronqué ( $\Delta$  *phoA*) et le gène *luc* forment un opéron synthétique.

1B. Séquence de la jonction entre *phoA* et *luc*.

Figure 2 :

- 5 Activités Luc et PhoA de *M. smegmatis* recombinant contenant le pJVED avec différents fragments nucléotidiques comme décrits en exemple.

Figure 3 :

- 10 Hybridation génomique (Southern blot) de l'ADN génomique de différentes espèces mycobactériennes à l'aide de la sonde correspondant aux fragments XXXXX de la séquence ID.

Figure 4 :

- 15 Représentation de l'hydrophobicité (Kyte et Doolittle) de la séquence codante du DP428 avec sa représentation schématique. Le motif LPISG précède immédiatement la région C-terminale hydrophobe. La séquence se termine par deux arginines.

Figure 5 :

- 20 Illustre la séquence SEQ ID N°1 correspondant à l'insert du vecteur pDP428 (déposé à la CNCM sous le N° I-1818).

Figure 6 :

- 25 Illustre la séquence SEQ ID N°2 correspondant à la région incluant le gène codant pour le DP428 (région soulignée). La figure fait apparaître que le DP428 fait probablement partie d'un opéron comprenant au moins trois gènes. La région doublement encadrée inclut probablement les régions promotrices.

- 30 La région simplement encadrée correspond au motif LPISG rappelant le motif LPXTG décrit chez les Gram positifs comme permettant l'ancrage aux peptidogluconnes.

Figure 7 :

Illustre la séquence SEQ ID N°3 correspondant à l'insert du vecteur p6D7 (déposé à la CNCM sous le N° I-1814).

A, B et C représentent les trois phases de lecture possible.

- 35 Figure 8 :

Illustre la séquence SEQ ID N°4 correspondant à l'insert du vecteur p5A3 (déposé à la CNCM sous le N° I-1815).

A,B et C représentent les trois phases de lecture possible.

5 Figure 9 :

Illustre la séquence SEQ ID N°5 correspondant à l'insert du vecteur p5F6 (déposé à la CNCM sous le N° I-1816).

A,B et C représentent les trois phases de lecture possible.

10 Figure 10 :

Illustre la séquence SEQ ID N°6 correspondant à l'insert du vecteur p2A29 (déposé à la CNCM sous le N° I-1817).

A,B et C représentent les trois phases de lecture possible.

15 Figure 11 :

Illustre la séquence SEQ ID N°7 correspondant à l'insert du vecteur p5B5 (déposé à la CNCM sous le N° I-1819).

A,B et C représentent les trois phases de lecture possible.

20 Figure 12 :

Illustre la séquence SEQ ID N°8 correspondant à l'insert du vecteur p1C7 (déposé à la CNCM sous le N° I-1820).

A,B et C représentent les trois phases de lecture possible.

25 Figure 13 :

Illustre la séquence SEQ ID N°9 correspondant à l'insert du vecteur p2D7 (déposé à la CNCM sous le N° I-1821).

A,B et C représentent les trois phases de lecture possible.

30 Figure 14 :

Illustre la séquence SEQ ID N°10 correspondant à l'insert du vecteur p1B7 (déposé à la CNCM sous le N° I-1843).

A,B et C représentent les trois phases de lecture possible.

35 Figure 15 :

Illustre la séquence SEQ ID N°11.

A,B et C représentent les trois phases de lecture possible.

Figure 16 :

Illustre la séquence SEQ ID N°12.

5 A,B et C représentent les trois phases de lecture possible.

Figure 17 :

Illustre la séquence SEQ ID N°13.

10 A,B et C représentent les trois phases de lecture possible.

Figure 18 :

Illustre la séquence SEQ ID N°14.

A,B et C représentent les trois phases de lecture possible.

15 Figure 19 :

Illustre la séquence SEQ ID N°15.

A,B et C représentent les trois phases de lecture possible.

Figure 20 :

20 Illustre la séquence SEQ ID N°16.

A,B et C représentent les trois phases de lecture possible.

Figure 21 :

Illustre la séquence SEQ ID N°17.

25 A,B et C représentent les trois phases de lecture possible.

Figure 22 :

Illustre la séquence SEQ ID N°18.

30 A,B et C représentent les trois phases de lecture possible.

Figure 23 :

Illustre la séquence SEQ ID N°19.

A,B et C représentent les trois phases de lecture possible.

35 Figure 24 :

Illustre la séquence SEQ ID N°20.

A,B et C représentent les trois phases de lecture possible.

Figure 25 :

Illustre la séquence SEQ ID N°21.

5 A,B et C représentent les trois phases de lecture possible.

Figure 26 :

Illustre la séquence SEQ ID N°22.

10 A,B et C représentent les trois phases de lecture possible.

Figure 27 :

Illustre la séquence SEQ ID N°23.

A,B et C représentent les trois phases de lecture possible.

15 Figure 28 :

Illustre la séquence SEQ ID N°24.

A,B et C représentent les trois phases de lecture possible.

Figure 29 :

20 Illustre les séquences SEQ ID N°25 et SEQ ID N°26 représentant un couple d'amorces utilisées pour amplifier spécifiquement par PCR la région correspondant aux nucléotides 964 à 1234 inclus dans la séquence SEQ ID N°1.

25

## EXEMPLES

### Matériel et méthodes

30

Cultures bactériennes, plasmides et milieux de cultures

Les cultures bactériennes et plasmides utilisées dans ces exemples sont regroupées dans le tableau 1. *E. coli* a été cultivé sur milieu liquide ou solide Luria-Bertani (LB). *M. smegmatis* a été cultivé sur milieu liquide Middlebrook 7H9 (Difco) additionné de dextrose albumine (ADC), 0,2 % de

35

glycérole et 0,05 % de Tween, ou sur milieu solide L. Si nécessaire, l'antibiotique kanamycine a été rajouté à une concentration de  $20 \mu\text{g/ml}^{-1}$ . Les clones bactériens présentant une activité PhoA ont été détectés sur de l'agar LB contenant  
5 du 5-bromo-4-chloro-3-indolye phosphate (X-P, à  $40 \mu\text{g/ml}^{-1}$ ).

#### Manipulation d'ADN et séquençage

Les manipulations d'ADN et les analyses par Southern  
10 blot ont été effectuées en utilisant les techniques standard (Sambrook, 1989). Les séquences d'ADN double brin ont été déterminées avec un kit de séquençage Taq Dye Deoxy Terminator Cycle (Applied Biosystems), dans un Système 9600 GeneAmp PCR (Perkin-Elmer), et après migration sur un système d'analyse ADN  
15 modèle 373 (Applied Biosystems).

#### Constructions des plasmides

Le plasmide pJVEDa a été construit à partir de pLA71,  
20 plasmide de transfert comportant le gène *phoA* tronqué et placé en phase avec *BlaF*. pLA71 a été coupé avec les enzymes de restriction *KpnI* et *NotI*, retirant ainsi *phoA* sans toucher le promoteur de *BlaF*. Le gène *luc* codant pour la luciférase de luciole a été amplifié à partir de pGEM-luc et un site de  
25 liaison du ribosome a été rajouté. *phoA* a été amplifié à partir de pJEM11. Les fragments amplifiés ont été coupés avec *PstI* et ligaturés ensemble. Les oligodéoxynucléotides utilisés sont les suivants :

pPV.luc.Fw : 5'GACTGCTGCAGAAGGAGAAGATCCAAATGG3'

30 luc.Bw : 5'GACTAGCGGCCGCGAATTCGTGACCTCCGAGG3'

pJEM.phoA.Fw : 5'CCGCGGATCCGGATACGTAC3'

phoA.Bw: 5'GACTGCTGCAGTTTATTTTCAGCCCCAGAGCG3'.

Le fragment ainsi obtenu a été réamplifié en utilisant les oligonucléotides complémentaires de ses extrémités, coupé  
35 avec *KpnI* et *NotI*, et intégré dans pLA71 coupé avec les mêmes

enzymes. La construction résultante a été électroporée dans *E. coli* DH5 $\alpha$  et *M. smegmatis* mc2 155. Un clone *M. smegmatis* émettant de la lumière et présentant une activité *phoA* a été sélectionné et appelé pJVED/*blaF*. L'insert a été retiré en utilisant *Bam*HI et la construction refermée sur elle-même, reconstruisant ainsi le pJVED<sub>a</sub>. Afin d'obtenir le pJVED<sub>b,c</sub>, le multisite de clonage a été coupé avec *Sca*I et *Kpn*I et refermé en enlevant un (pJVED<sub>b</sub>) ou deux (pJVED<sub>c</sub>) nucléotides du site *Sna*BI. Après fusion six cadres de lecture ont pu ainsi être obtenus. L'insert du pJVED/*hsp18* a été obtenu par amplification en chaîne par polymérase (ACP) de pPM1745 (Servant, 1995) en utilisant des oligonucléotides de la séquence :

18.Fw : 5'GTACCAGTACTGATCACCCGTCTCCCGCAC3'

18.Back : AGTCAGGTACCTCGCGGAAGGGGTCAGTGCG3'

Le produit a été coupé avec *Kpn*I et *Sca*I, et ligaturé à pJVED<sub>a</sub>, coupé avec les mêmes enzymes, quittant ainsi le pJVED/*hsp18*.

Le pJVED/*P19kDa* et le pJVED/*erp* furent construits en coupant avec *Bam*HI l'insert de pExp410 et pExp53 respectivement, et en les insérant dans le site *Bam*HI du multisite de clonage de pJVED<sub>a</sub>.

#### Mesure de l'activité phosphatase alcaline

*M. smegmatis* ont été cultivés dans un milieu LB additionnés de 0,05 % de Tween 80 (Aldrich) et de kanamycine (20  $\mu$ g/ml<sup>-1</sup>) à 37°C pendant 24 heures. L'activité de la phosphatase alcaline a été mesurée par la méthode de Brockman et Heppel (Brockman, 1968) dans un extrait soniqué, avec *p*-nitrophénylphosphate comme substrat de la réaction. La quantité de protéines a été mesurée par essai Bio-Rad. L'activité phosphatase alcaline est exprimée en unité arbitraire (densité optique à 420 nm x  $\mu$ g de protéines<sup>-1</sup> x minutes<sup>-1</sup>).

### Mesure de l'activité luciférase

*M. smegmatis* a été cultivé dans un milieu LB additionné  
5 de 0,05 % de Tween 80 (Aldrich) et de kanamycine ( $20 \mu\text{g/ml}^{-1}$ )  
à 37°C pendant 24 heures et utilisé en pleine croissance  
exponentielle (DO à 600 nm comprise entre 0,3 et 0,8). Les  
aliquots de suspensions bactériennes ont été brièvement  
soniqués et l'extrait cellulaire a été utilisé pour mesurer  
10 l'activité de la luciférase. 25  $\mu\text{l}$  de l'extrait soniqué ont été  
mélangés avec 100  $\mu\text{l}$  de substrat (système d'essai luciférase  
Promega) automatiquement dans un luminomètre et la lumière  
émise exprimée en ULR (Unités Lumineuses Relatives). Les  
bactéries ont été comptées par dilutions sérielles de la  
15 suspension d'origine sur milieu agar LB kanamycine et  
l'activité de la luciférase exprimée en ULR/ $\mu\text{g}$  de protéines  
bactériennes ou en ULR/ $10^3$  bactéries.

### Construction de banques génomiques de *M. tuberculosis*

20

Les banques ont été obtenues en utilisant  
essentiellement pJVED<sub>a,b,c</sub> précédemment décrit.

Préparation de macrophages issus de la moelle osseuse et  
25 infection par *M. smegmatis* recombinants

Les macrophages issus de la moelle osseuse ont été  
préparés comme décrits par Lang, 1991. En résumé, les cellules  
de la moelle osseuse ont été prélevés du fémur de souris  
30 C57BL/6 âgée de 6 à 12 semaines (Iffa-Credo, France). Les  
cellules en suspensions ont été lavées et resuspendues dans du  
DMEM enrichi avec 10 % de sérum foetal de veau, 10 % de milieu  
L-cell conditionné et 2 mM de glutamine, sans antibiotiques.  
10<sup>6</sup> cellules ont étéensemencées sur des plaques 24 puits  
35 Costar à fond plat dans 1 ml. Après quatre jours à 37°C dans  
une atmosphère humide à 10 % de teneur en CO<sub>2</sub>, les macrophages  
ont été rincés et réincubés pendant deux à quatre jours



supplémentaires. Les cellules d'un puits contrôle ont été lysées avec du triton x 100 à 0,1 % dans l'eau et les noyaux énumérés. Environ  $5 \times 10^5$  cellules adhérentes ont été comptées. Pour l'infection, *M. smegmatis* portant les différents plasmides a été cultivé en pleine phase exponentielle ( $DO_{600nm}$  entre 0,4 et 0,8) et dilué jusqu'à une DO de 0,1 puis 10 fois dans un milieu pour macrophage. 1 ml a été ajouté à chaque puits et les plaques ont été centrifugées et incubées quatre heures à 37°C. Après trois lavages, les cellules ont été incubées dans un milieu contenant de l'amykacine pendant deux heures. Après trois nouveaux lavages, les cellules infectées adhérentes ont été incubées dans un milieu macrophage pendant une nuit. Les cellules ont ensuite été lysées dans 0,5 ml de tampon de lyse (Promega). 100  $\mu$ l ont été soniqués et la lumière émise a été mesurée sur 25  $\mu$ m. Simultanément, les bactéries ont été énumérées par étalement sur L-agar-kanamycine (20  $\mu$ g/ml<sup>-1</sup>). La lumière émise est exprimée en ULR/10<sup>3</sup> bactéries.

#### Analyses des banques de données

Les séquences nucléotidiques ont été comparées à EMBL et GenBank en utilisant l'algorithme FASTA et les séquences protéiques ont été analysées par similitude grâce aux banques de données PIR et Swiss Prot en utilisant l'algorithme BLAST.

#### Exemple 1 : Les vecteurs pJVED

Les vecteurs pJVED (Figure 1A) sont des plasmides portant un gène *phoA* tronqué de *E. coli* dépourvu de codon d'initiation, de séquence signal et de séquence régulatrice. Le site multiple de clonage (SMC) permet l'insertion de fragments des gènes codants pour d'éventuelles protéines exportées ainsi que leurs séquences de régulation. Dès lors, la protéine de fusion peut être produite et présenter une activité phosphatase alcaline si elle est exportée. Seules les fusions en phase pourront être productives. Ainsi, le SMC a

été modifié de sorte que les fusions peuvent être obtenues dans six phases de lecture. En aval de *phoA*, le gène *luc* de la luciférase de luciole a été inséré. Le gène complet avec le codon d'initiation mais sans qu'aucun promoteur n'ait été  
5 utilisé devrait ainsi s'exprimer avec *phoA* comme dans un opéron synthétique. Un nouveau site de liaison des ribosomes a été inséré huit nucléotides en amont du codon d'initiation de *luc*. Deux terminateurs transcriptionnels sont présents dans les vecteurs pJVED, un en amont du SMC et un second en aval de *luc*.  
10 Ces vecteurs sont des plasmides de transfert *E. coli*-*mycobacterium* avec un gène de résistance à la kanamycine comme marqueur de sélection.

*phoA* et *luc* fonctionnent comme dans un opéron, mais l'exportation est nécessaire pour l'activité *phoA*.

15 Quatre plasmides ont été construits par insertion dans le SMC de fragments d'ADN d'origine diverse :

Dans la première construction nommée pJVED/*blaF*, le fragment de 1,4 kb provient du plasmide déjà décrit pLA71 (Lim, 1995). Ce fragment issu du gène  $\beta$ -lactamase (*blaF*) de *M.*  
20 *fortuitum* D216 (Timm, 1994) inclut le promoteur muté hyperactif, le segment codant pour 32 acides aminés de la séquence signal et les 5 premiers acides aminés de la protéine mature. Ainsi cette construction inclut le promoteur le plus fort connu chez *mycobacterium* et les éléments nécessaires à  
25 l'exportation de la protéine de la fusion *phoA*. Par conséquent, on peut attendre de cette construction une forte émission de lumière et une bonne activité *phoA*.

Dans une deuxième construction nommée pJVED/*hsp18*, un fragment de 1,5 kb a été cloné à partir du plasmide déjà décrit  
30 pPM1745 (Servant, 1995). Ce fragment inclut les nucléotides codants pour les dix premiers acides aminés de la protéine de choc thermique de 18 kb issue de *Streptomyces albus* (heat shock protein 18, HSP 18), le site de liaison du ribosome, le promoteur et, en amont, des sites régulateurs contrôlant son

expression. Cette protéine appartient à la famille de alpha-crystalline de HSP à faible poids moléculaire (Verbon, 1992). Son homologue issu de *M. leprae*, l'antigène de 18 kDa, est déjà connu pour être induit durant la phagocytose par un macrophage murin de la lignée cellulaire J-774 (Dellagostin, 1995). Dans des conditions de culture standard, le pJVED/hsp18, montre une faible activité *luc* et aucune activité *phoA*.

Dans une troisième construction, nommée pJVED/P19kDa, l'insert issu de pExp410 (Lim, 1995) a été coupé et cloné dans le SMC de pJVED<sub>a</sub>. Ce fragment inclut les nucléotides codants pour les 134 premiers acides aminés de la protéine connue de *M. tuberculosis* 19 kDa et de ses séquences régulatrices. Comme cela a pu être mis en évidence, cette protéine est une lipoprotéine glycosilée (Garbe, 1993 ; Herrmann, 1996). Une partie de cette lipoprotéine a été fusionnée avec une protéine A de surface issue de *Borrelia burgdorferi* pour construire un vaccin recombinant de *M. bovis* BCG capable d'induire une forte réponse immunitaire (Stover, 1993). Sur la figure 2, on observe, pour cette construction, une bonne activité *luc* correspondant à un promoteur fort, mais l'activité *phoA* est, de loin, la plus forte des quatre constructions. L'activité *phoA* élevée de cette protéine de fusion avec une lipoprotéine s'explique par le fait qu'elle reste attachée à la paroi cellulaire par son extrémité N-terminal.

Dans la quatrième et dernière construction nommée pJVED/erp l'insert provient de pExp53 (Lim, 1995) et a été cloné dans le SMC de pJVED<sub>a</sub>. pExp53 est le plasmide initial sélectionné pour son activité *phoA* et contenant une partie du gène *erp* de *M. tuberculosis* qui code pour un antigène de 28 kDa. Ce dernier inclut la séquence signal, une partie de la protéine mature et, en amont du codon d'initiation, le site de liaison de ribosome. Le promoteur a été cartographié. Une boîte fer (iron box) putative du type *fur* est présente dans cette région et encadre la région -35 du promoteur (Berthet, 1995).

Comme prévu (figure 2) cette construction présente une bonne émission lumineuse et une bonne activité *phoA*. Le fait que cette protéine de fusion, contrairement à la fusion avec la lipoprotéine de 19 kDa, ne semble pas attachée à la paroi cellulaire n'exclut pas que la protéine native y soit associée. De plus, l'extrémité C-terminal de *erp* est absente de la protéine de fusion.

**Exemple 2 :** Construction d'une banque d'ADN génomique de *M. tuberculosis* dans les vecteurs pJVED<sub>g</sub> et identification d'un des membres de ces banques, (DP428), induit au cours de la phagocytose par les macrophages murins dérivés de la moelle osseuse.

Les différentes constructions sont testées pour leur capacité à évaluer l'expression intracellulaire des gènes identifiés par l'expression de *phoA*. Dans cet objectif, l'activité *luc* est exprimée en URL pour  $10^3$  bactéries en culture axénique dans des conditions intracellulaires. L'induction ou la répression suivant la phagocytose par les macrophages murins dérivés de la moelle osseuse peut être évaluée convenablement par la mesure des activités spécifiques. Les résultats de deux expériences distinctes sont présentés dans le tableau 2.

Le plasmide pJVED/*hsp18* a été utilisé comme contrôle positif pour l'induction durant la phase de croissance intracellulaire. Malgré que l'induction du promoteur par le chauffage de la bactérie à 42°C n'ait pas été concluant la phagocytose de la bactérie conduit clairement à une augmentation de l'activité du promoteur. Dans toutes les expériences, l'activité *luc* intracellulaire a été fortement induite, augmentant de 20 à 100 fois l'activité basale initialement faible (Servant, 1995).

Le plasmide pJVED/*blaF* a été utilisé comme contrôle de la modulation non spécifique au cours de la phagocytose. De faibles variations ont pu être mises en évidence, probablement dues à des changements de conditions de cultures. Quoi qu'il en

soit, ces faibles variations ne sont pas comparables à l'induction observée avec le plasmide pJVED/hsp18.

Tous les membres de la banque d'ADN ont été testés par mesure de l'activité du promoteur durant la croissance  
5 intracellulaire. Parmi eux, le DP428 est fortement induit au cours de la phagocytose (tableau 1).

TABLEAU 1 (cf. après les exemples).

10 **Exemple 3 :** La séquence complète du gène DP428 et de ses régions flanquantes.

Une sonde de la région codante de DP428 a été obtenue par ACP, et utilisée pour hybrider l'ADN génomique de différentes espèces de mycobactéries. D'après les résultats de  
15 la figure 3, le gène est présent uniquement dans les mycobactéries du complexe de *M. tuberculosis*.

L'analyse de la séquence montre que DP428 fait partie d'un opéron. La séquence codante et les régions flanquantes ne présentent aucune homologie avec des séquences connues déposées  
20 dans les banques de données.

D'après la séquence codante, ce gène code pour une protéine de 10 kDa avec un peptide signal, une extrémité C-terminal hydrophobe terminée par deux arginines et précédée par un motif LPISG semblable au motif connu LPXTG. Ces deux  
25 arginines pourraient correspondre à un signal de rétention et la protéine DP428 pourrait être accrochée par ce motif à des peptidoglycanes comme cela a déjà été décrit chez d'autres bactéries Gram+.

Le mécanisme de survie et de croissance intracellulaire  
30 des mycobactéries est complexe et les relations intimes entre la bactérie et la cellule hôte restent inexpliquées. Quel que soit le mécanisme, la croissance et la survie intracellulaire des mycobactéries dépend de facteurs produits par la bactérie et capables de moduler la réponse de l'hôte. Ces facteurs  
35 peuvent être des molécules exposées à la surface cellulaire telle que LAM ou des protéines associées à la surface cellulaire, ou des molécules activement secrétées.

D'un autre côté, intracellulairement, les bactéries elles-mêmes doivent faire face à un environnement hostile. Elles semblent y répondre par des moyens proches de ceux mis en oeuvre dans les conditions de stress, par l'induction de protéines de choc thermique (Dellagostin, 1995), mais aussi par induction ou la répression de différentes protéines (Lee, 1995). En utilisant une méthodologie dérivée de la PCR, Plum et Clark-curtiss (Plum, 1994) ont montré qu'un gène de *M. avium* inclu dans un fragment d'ADN de 3 kb, est induit après la phagocytose par des macrophages humains. Ce gène code pour une protéine putative exportée comprenant une séquence leader mais ne présentant pas d'homologie significative avec les séquences proposées par les banques de données. L'induction, pendant la phase de croissance intracellulaire, d'une protéine de choc thermique de faible poids moléculaire issue de *M. leprae* a également été mise en évidence (Dellagostin, 1995). Dans une autre étude, les protéines bactériennes de *M. tuberculosis* ont été métaboliquement marquées pendant la phase de croissance intracellulaire ou bien dans des conditions de stress et séparées par électrophorèse sur gel à deux dimensions : 16 protéines de *M. tuberculosis* ont été induites et 28 réprimées. Les mêmes protéines sont mises en jeu au cours de stress provoqué par un faible pH, un choc thermique, H<sub>2</sub>O<sub>2</sub>, ou au cours de la phagocytose par des monocytes humains de la lignée THP1. Quoi qu'il en soit, le comportement des protéines induites et réprimées était unique dans chaque condition (Lee, 1995). Pris ensemble, ces résultats indiquent qu'un dialogue moléculaire subtile est mis en place entre les bactéries et leurs hôtes cellulaires. De ce dialogue dépend probablement le sort de l'organisme intracellulaire.

Dans ce contexte, l'induction de l'expression de DP428 pourrait être d'une importance majeure, indiquant un rôle important de cette protéine dans la survie et la croissance intracellulaire.

La méthode utilisée dans ces expériences pour évaluer

l'expression intracellulaire des gènes présente l'avantage d'être simple comparée aux autres techniques comme la technique décrite par Mahan et al. (Mahan, 1993) adaptée aux mycobactéries et proposée par Bange et al. (Bange, 1996), ou la méthode substructive basée sur l'ACP décrite par Plum et Clark-curtiss (Plum, 1994). Il existe indiscutablement une variabilité comme le montre la comparaison des différentes expériences. Bien que provoquer l'induction ou la répression soit suffisant il est désormais possible de l'évaluer fournissant ainsi un outil supplémentaire d'études physiologiques des protéines exportées identifiées par fusion avec *phoA*.

#### Exemple 4 :

Recherche d'une modulation de l'activité des promoteurs lors des phases intramacrophagiques

Des macrophages de moelle osseuses de souris sont préparés comme décrit par Lang et Antoine (Réf. 13). Les bactéries de *M. segmentis* recombinantes, dont on a déterminé l'activité luciférase par  $10^3$  bactéries comme précédemment, sont incubées à 37°C sous atmosphère humidifiée et enrichie en CO<sub>2</sub> à 5%, pendant 4 heures en présence de ces macrophages de telle manière qu'elles soient phagocytées. Après rinçage pour éliminer les bactéries extracellulaires restantes, on ajoute au milieu de culture de l'amikacine (100 µg/ml) pendant deux heures. Après un nouveau rinçage, le milieu est remplacé par un milieu de culture (DMEM enrichi de 10% de sérum de veau et 2mM de glutamine) sans antibiotiques. Après une nuit d'incubation comme précédemment, les macrophages sont lysés à froid (4°C) à l'aide d'un tampon de lyse (cell lysis buffer, Promega), et l'activité luciférase par  $10^3$  bactéries déterminée. Le rapport des activités à la mise en culture et après une nuit donne le coefficient d'induction.

Tableau 1

construct	% recovery	RLU/10 <sup>3</sup> bacteria extracellular	RLU/10 <sup>3</sup> bacteria intracellular	induction
pJVED/blaF*	0.5	1460	1727	1.2
pJVED/hsp18	0.6	8	57	7.1
pJVED/DP428	0.7	0.06	18	300
construct	% recovery C57BL/6 Balb/C	RLU/10 <sup>3</sup> bacteria extracellular	RLU/10 <sup>3</sup> bacteria intracellular C57BL/6 Balb/C	induction C57BL/6 Balb/C
pJVED/blaF*	7	662	250	0.4
pJVED/hsp18	6.7	164	261	1.4
pJVED/DP428	1.6	0.08	1.25	1.6
			3.3	2
				41



## 5 Références bibliographiques

- AIDS therapies, 1993, in Mycobacterial infections, ISBN 0-9631698-1-5, pp. 1-11.
- Barany F., 1911, Proc. Natl. Acad. Sci. USA, 88:189-193.
- Bates J. et al., 1986, Am. Rev. Respir. Dis., 134:415-417.
- 10 Bates, J. 1979. Chest. 76(Suppl.):757-763.
- Bates, J. et al.. 1986. Am. Rev. Respir. Dis. 134:415-417.
- Bourgoin, A., G. Agius. 1994. Rev. Fr. Lab. 273:21-26.
- Bouvet, E. 1994. Rev. Fr. Lab. 273:53-56.
- Burg J.L. et al., 1996, Mol. and Cell. Probes, 10:257-271.
- 15 Chevrier D. et al., 1993, Mol. and Cell. Probes, 7:187-197.
- Chu B.C.F. et al., 1986, Nucleic Acids Res., 14:5591-5603.
- Daniel, T.M. et al. 1987. Am. Rev. Respir. Dis. 135:1137-1151).
- Drake, T.A. et al. 1987. J. Clin. Microbiol. 25:1442-1445.
- Duck P. et al., 1990, Biotechniques, 9:142-147.
- 20 Dunn A.R. et al., 1977, Cell. 12:23.
- Erlich, H.A. 1989. In PCR Technology. Principles and Applications for DNA Amplification. New York: Stockton Press.
- Felgner et al., 1987, Proc. Natl. Acad. Sci., 84:7413.
- Fraley et al., 1980, J. Biol. Chem., 255:10431.
- 25 Guateli J.C. et al., 1990, Proc. Natl. Acad. Sci. USA, 87:1874-1878.
- Houbenweyl, 1974, in Meuthode der Organischen Chemie, E. Wunsch Ed., Volume 15-I et 15-II, Thieme, Stuttgart.
- Huygen K. et al., 1996, Nature Medicine, 2(8):893-898.
- 30 Innis, M.A. et al. 1990. in PCR Protocols. A guide to Methods and Applications. San Diego: Academic Press.
- Kaneda et al., 1989, Science, 243:375

- Kent P.T. et al., 1985, Public Health Mycobacteriology : a Guide for Level III Laboratory. Center for Disease Control Ed., Atlanta, USA.
- Kiehn, T.E., et al. 1987. J. Clin. Microbiol. 25:1551-1552.
- 5 Kievitits T. et al., 1991, J. Virol. Methods, 35:273-286.
- Kohler G. et al., 1975, Nature, 256(5517):495-497.
- Kwoh D.Y. et al., 1989, Proc. Natl. Acad. Sci. USA, 86:1173-1177.
- Landegren U. et al., 1988, Science, 241:1077-1080.
- 10 Lizardi P.M. et al., 1988, Bio/technology, 6:1197-1202.
- Matthews J.A. et al., 1988, Anal. Biochem., 169:1-25.
- Merrifield R.D., 1966, J. Am. Chem. Soc., 88(21):5051-5052.
- Midoux, 1993, Nucleic Acids Research, 21:871-878/
- Miele E.A. et al., 1983, J. Mol. Biol., 171:281-295.
- 15 Milstein C., 1981, Annu. Rev. Biochem., 50:657-680.
- Milstein C., 1981, Sci. Am., 243(4):66-74.
- Minton N.P., 1984, Gene, 31, 269-273.
- Montgomery et al., 1993, DNA Cell Biol., 12:777-783.
- Pagano et al., 1967, J. Virol., 1:891
- 20 Palva A. et al., 1984, FEMS Microbiol. Lett., 23:83.
- Pastore, 1994, Circulation, 90:I-517.
- Patel et al. 1990. J. Clin. Microbiol. 28:513-518.
- Pelacic V. et al., 1996, FEMS Microbiol. Letters, 144:161-166.
- Pelacic V. et al., 1996, J. Bact., 178(4):1197-1199.
- 25 Polsky-Cynkin R. et al., 1985, J. Clin. Chem., 31:1438.
- Ranki M. et al., 1983, Gene, 21:77.
- Roberts, M.C., et al. 1987. J. Clin. Microbiol. 25:1239-1243.
- Rolfs, A. et al. 1991. In PCR Topics. Usage of Polymerase Chain reaction in Genetic and Infectious Disease. Berlin: Springer-
- 30 Verlag.
- Saiki R.K. et al., 1988, Science, 239:487-491.
- Saiki, R.K. et al. 1985. Science. 230:1350-1354.

- Sambrook, J. et al. 1989. In Molecular cloning : A Laboratory Manual. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press.
- Sanchez-Pescador R., 1988, J. Clin. Microbiol., 26(10):1934-1938.
- Segev D., 1992, in « Non-radioactive Labeling and Detection of Biomolecules ». Kessler C. Springer Verlag, Berlin, New-York, 197-205.
- Shinnick T.M., 1987, J. Bacteriol., 169:1080-1088.
- Shiver J.W., 1995, in Vaccines 1995, eds Chanock, R.M. Brown, F. Ginsberg, H.S. & Norrby, E.), pp.95-98, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Spargo C.A. et al., 1996, Mol. and Cell. Probes, 10:247-256
- Stone B.B. et al., 1996, Mol. and Cell. Probes, 10:359-370.
- Tascon R.E et al., 1996, Nature Medicine, 2(8):888-892.
- Technique assemblage oligonucléotides, 1983, Proc. Natl. Acad. Sci. USA, 80:7461-7465.
- Technique des bêta-cyanethylphosphoramidites, 1986, Bioorganic Chem., 4:274-325.
- Thierry D. et al. 1990. Nucl. Acid Res. 18:188.
- Tuberculosis Prevention Trial, 1980, Mendis, « Trial of BCG vaccines in South India for Tuberculosis Infection », Indian Journal of Medical research, 1972 (Suppl.):1-74.
- Urdea M.S. et al., 1991, Nucleic Acids Symp. Ser., 24:197-200.
- Urdea M.S., 1988, Nucleic Acids Research, 11: 4937-4957.
- Vega-Lopez F. et al., 1993, Infect. Immun., 61(5):2145-2153.
- Walker G.T. et al., 1992, Nucleic Acids Res., 20:1691-1696.
- Walker G.T. et al., 1992, Proc. Natl. Acad. Sci. USA, 89:392-396.
- Yuen L.K.W. et al., 1993, J. Clin. Microbiol., 31, 1615-1618.

1. Bange, F.C., A.M. Brown, and W.R. Jacobs JR. 1996. Leucine auxotrophy restricts growth of *Mycobacterium bovis* BCG in macrophages. *Infect. Immun.* 64: 1794-1799.
2. Berthet, F.X., J. Rauzier, E.M. Lim, W. Philipp, B. Gicquel, and D. Portnoï. 1995. Characterization of the *M. tuberculosis* *erp* gene encoding a potential cell surface protein with repetitive structures. *Microbiology*. In press:
3. Boquet, P.L., C. Manoil, and J. Beckwith. 1987. Use of *TnphoA* to detect genes for exported proteins in *Escherichia coli*: identification of the plasmid-encoded gene for a periplasmic phosphatase. *J. Bacteriol.* 169: 1663-1669.
4. Brockman, R.W. and L.A. Heppel. 1968. On the localization of alkaline phosphatase and cyclic phosphodiesterase in *Escherichia coli*. *Biochemistry*. 7: 2554-2561.
5. Clemens, D.L. 1996. Characterization of the *Mycobacterium tuberculosis* phagosome. *Trends Microbiol.* 4: 113-118.
6. Clemens, D.L. and M.A. Horwitz. 1995. Characterization of the *Mycobacterium tuberculosis* phagosome and evidence that phagosomal maturation is inhibited. *J. Exp. Med.* 181: 257-270.
7. Dellagostin, O.A., G. Esposito, L.-J. Eales, J.W. Dale, and J.J. McFadden. 1995. Activity of mycobacterial promoters during intracellular and extracellular growth. *Microbiol.* 141: 2123-2130.
8. Gaillard, J.L., P. Berche, C. Frehel, E. Gouin, and P. Cossart. 1991. Entry of *L. monocytogenes* into cells is mediated by internalin, a repeat protein reminiscent of surface antigens from Gram-positive cocci. *Cell*. 65: 1127-1141.

9. Garbe, T., D. Harris, M. Vordermeir, R. Lathigra, J. Ivanyi, and D. Young. 1993. Expression of the *Mycobacterium tuberculosis* 19-kilodalton antigen in *Mycobacterium smegmatis*: immunological analysis and evidence of glycosylation. Inf. Immun. 61: 260-267.
10. Herrmann, J.L., P. O'Gaora, A. Gallagher, J.E.R. Thole, and D.B. Young. 1996. Bacterial glycoproteins: a link between glycosylation and proteolytic cleavage of a 19 kDa antigen from *Mycobacterium tuberculosis*. EMBO J. 15: 3547-3554.
- 10 11. Hoffman, C.S. and A. Wright. 1985. Fusion of secreted proteins to alkaline phosphatase: an approach for studying protein secretion. Proc. Natl. Acad. Sci. USA. 82: 5107-5111.
12. Isberg, R.R., D.L. Voorhis, and S. Falkow. 1987. Identification of invasins: a protein that allows enteric bacteria to penetrate cultured mammalian cells. Cell. 50: 769-778.
13. Lang, T. and J.-C. Antoine. 1991. Localization of MHC classII molecules in murine bone marrow-derived macrophages. Immunology. 72: 199-205.
- 20 14. Lee, B.Y. and M.A. Horwitz. 1995. Identification of macrophage and stress-induced proteins of *Mycobacterium tuberculosis*. J. Clin. Invest. 96: 245-249.
15. Lim, E.M., J. Rauzier, J. Timm, G. Torrea, A. Murray, B. Gicquel, and D. Portnoï. 1995. Identification of *Mycobacterium tuberculosis* DNA sequences encoding exported proteins, using *phoA* gene fusions. J. Bacteriol. 177: 59-65.
- 25

16. Mahan, M.J., J.M. Slauch, and M. J.J. 1993. Selection of bacterial virulence genes that are specifically induced in host tissues. *Science*. 259: 686-688.
17. Pelicic, V., J.M. Reyrat, and B. Gicquel. 1996. 5 Generation of unmarked directed mutations in mycobacteria, using sucrose counterselectable suicide vectors. *Mol. Microbiol.* In press.:
18. Pettersson, J., R. Nordfelth, E. Dubinina, T. Bergman, M. Gustafsson, K.E. Magnusson, and H. Wolf-Watz. 1996. 10 Modulation of virulence factor expression by pathogen target cell contact. *Science*. 273: 1231-1233.
19. Plum, G. and J.E. Clark-Curtiss. 1994. Induction of *Mycobacterium avium* gene expression following phagocytosis by human macrophages. *Infect. Immun.* 62: 476-483.
- 15 20. Sambrook, J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning. A laboratory manual. 2nd ed., Vol. Cold Spring Harbor. New York : Cold Spring Harbor Laboratory Press.
21. Servant, P. and P. Mazodier. 1995. Characterization of *Streptomyces albus* 18-kilodalton heat shock-responsive 20 protein. *J. Bacteriol.* 177: 2998-3003.
22. Stover, C.K., G.P. Bansal, M.S. Hanson, S.R. Burlein, S.R. Palaszynski, J.F. Young, S. Koenig, D.B. Young, A. Sadziene, and A.G. Barbour. 1993. Protective immunity elicited by recombinant Bacille Calmette-Guerin (BCG) expressing outer 25 surface protein A (OspA) lipoprotein: a candidate Lyme disease vaccine. *J. Exp. Med.* 178: 197-209.
23. Sturgill-Koszycki, S., P.H. Schlesinger, P. Chakroborty, P.L. Haddix, H.L. Collins, A.K. Fok, R.D. Allen, S.L. Gluck,

- J. Heuser, and D.G. Russell. 1994. Lack of acidification in *Mycobacterium* phagosomes by exclusion of the vesicular proton-ATPase. *Science*. 263: 678-681.
24. Timm, J., M.G. Perilli, C. Duez, J. Trias, G. Orefici, L. Fattorini, G. Amicosante, A. Oratore, B. Boris, J.M. Frere, A.P. Pugsley, and B. and Gicquel. 1994. Transcription and expression analysis, using *lacZ* and *phoA* gene fusions, of *Mycobacterium fortuitum* B-lactamase genes cloned from a natural isolate and a high-level B-lactamase producer. *Mol. Microbiol.* 12: 491-504.
25. Verbon, A., R.A. Hartskeerl, A. Schuitema, A.H. Kolk, D.B. Young, and R. Lathigra. 1992. The 14,000-molecular-weight antigen of *Mycobacterium tuberculosis* is related to the alpha-crystallin family of low-molecular-weight heat shock proteins. *J Bacteriol.* 174: 1352-1359.
26. Xu, S., A. Cooper, S. Sturgill-Koszycki, T. van Heyningen, D. Chatterjee, I. Orme, P. Allen, and D.G. Russel. 1994. Intracellular trafficking in *Mycobacterium tuberculosis* and *Mycobacterium avium*-infected macrophages. *J. Immunol.* 153: 2568-2578.

## Revendications

1. Vecteur recombinant de criblage, de clonage et/ou d'expression, caractérisé en ce qu'il se réplique chez des  
5 mycobactéries et en ce qu'il contient :
  - 1) un réplicon fonctionnel chez les mycobactéries ;
  - 2) un marqueur de sélection ;
  - 3) une cassette reporteur comprenant :
    - a) un site de clonage multiple (polylinker),
    - 10 b) éventuellement un terminateur de transcription actif chez les mycobactéries, en amont du polylinker,
    - c) une séquence nucléotidique codante issue d'un gène codant pour un marqueur d'expression, d'exportation et/ou de sécrétion de protéine, ladite séquence nucléotidique étant  
15 dépourvue de son codon d'initiation et de ses séquences de régulation, et
    - d) une séquence nucléotidique codante issue d'un gène codant pour un marqueur d'activité de promoteurs contenus dans le même fragment, ladite séquence nucléotidique étant pourvue  
20 de son codon d'initiation.
2. Vecteur recombinant selon la revendication 1, caractérisé en ce que la séquence nucléotidique codante issue d'un gène codant pour un marqueur d'expression, d'exportation et/ou de sécrétion  
25 de protéine est une séquence codante issue du gène *phoA* de la phosphatase alcaline.
3. Vecteur recombinant selon l'une des revendications 1 et 2, caractérisé en ce que la séquence nucléotidique codante issue  
30 d'un gène codant pour un marqueur d'expression, d'exportation et/ou de sécrétion de protéine est une séquence codante du gène de la  $\beta$ -agarase, de la nucléase d'un staphylocoque ou de la  $\beta$ -lactamase d'une mycobactérie.
- 35 4. Vecteur recombinant selon l'une des revendications 1 à 3, caractérisé en ce que la séquence nucléotidique codante issue d'un gène codant pour un marqueur d'activité de promoteurs



contenus dans le même fragment est une séquence codante issue du gène *luc* de la luciférase de luciole.

5 5. Vecteur recombinant selon l'une des revendications 1 à 4, caractérisé en ce que la séquence nucléotidique codante issue d'un gène codant pour un marqueur d'activité de promoteurs contenus dans le même fragment est une séquence codante issue du gène *GFP* de la Green Fluorescent Protein.

10 6. Vecteur recombinant selon l'une des revendications 1 à 5, caractérisé en ce que le terminateur de transcription actif chez les mycobactéries est le terminateur du coliphage T4 (tT4).

15 7. Vecteur recombinant selon l'une des revendications 1 à 6, caractérisé en ce qu'il est un plasmide choisi parmi les plasmides suivants déposés à la CNCM (Collection Nationale de Cultures de Microorganismes, Paris, France) :

20 a) pJEVDa déposé à la CNCM sous le N° I-1797, le 12/12 1996,

b) pJEVDb déposé à la CNCM sous le N° I-1906, le 25 juillet 1997,

c) pJEVDC déposé à la CNCM sous le N° I-1799, le 12/12 1996,

25 d) pJEVD/*M. tuberculosis* déposé à la CNCM sous le N° I-1907, le 25 juillet 1997.

8. Vecteur recombinant selon l'une des revendications 1 à 7, caractérisé en ce qu'il comprend en l'un des sites de clonage  
30 du polylinker une séquence d'acide nucléique de mycobactérie chez laquelle on détecte un polypeptide susceptible d'être exporté et/ou sécrété, et/ou d'être induit ou réprimé lors de l'infection par ladite mycobactérie ou encore produit de façon constitutive, ainsi que les séquences promotrices et/ou  
35 régulatrices associées susceptibles de permettre ou de favoriser l'exportation et/ou la sécrétion dudit polypeptide, ou tout ou partie de gène codant pour ledit polypeptide.

9. Vecteur recombinant selon l'une des revendications 1 à 8, caractérisé en ce que la séquence d'acide nucléique de mycobactérie qu'il contient est obtenue par digestion enzymatique de l'ADN génomique ou de l'ADN complémentaire d'un  
5 ARN d'une mycobactérie pathogène.

10. Vecteur recombinant selon l'une des revendications 1 à 9, caractérisé en ce que ladite mycobactérie pathogène est *M. tuberculosis*.

10

11. Vecteur recombinant selon l'une des revendications 1 à 9, caractérisé en ce que ladite mycobactérie pathogène est choisie parmi *M. africanum*, *M. bovis*, *M. avium* ou *M. leprae*

15 12. Vecteur recombinant selon la revendications 10, caractérisé en ce qu'il est un plasmide choisi parmi les plasmides suivants déposés à la CNCM:

- a)p6D7 déposé le 28 janvier 1997 à la CNCM sous le N°I-1814,
- b)p5A3 déposé le 28 janvier 1997 à la CNCM sous le N°I-1815,
- 20 c)p5F6 déposé le 28 janvier 1997 à la CNCM sous le N°I-1816,
- d)p2A29 déposé le 28 janvier 1997 à la CNCM sous le N°I-1817,
- e)pDP428 déposé le 28 janvier 1997 à la CNCM sous le N°I-1818,
- f)p5B5 déposé le 28 janvier 1997 à la CNCM sous le N°I-1819,
- g)p1C7 déposé le 28 janvier 1997 à la CNCM sous le N°I-1820,
- 25 h)p2D7 déposé le 28 janvier 1997 à la CNCM sous le N°I-1821,
- i)p1B7 déposé le 31 janvier 1997 à la CNCM sous le N°I-1843.

13. Vecteur recombinant selon la revendication 12, caractérisé en ce qu'il s'agit du plasmide pDP428 déposé le 28 janvier 1997  
30 à la CNCM sous le N°I-1818.

14. Procédé de criblage de séquences de nucléotides issues de mycobactéries pour déterminer la présence de séquences correspondant à des polypeptides exportés et/ou sécrétés  
35 pouvant être induits ou réprimés lors de l'infection, leurs séquences promotrices et/ou régulatrices associées susceptibles notamment de permettre ou de favoriser

l'exportation et/ou la sécrétion desdits polypeptides d'intérêt, ou tout ou partie de gènes d'intérêt codant pour lesdits polypeptides, caractérisé en ce qu'il met en oeuvre un vecteur selon l'une des revendications 1 à 13.

5

15. Procédé de criblage selon la revendication 14, caractérisé en ce qu'il comprend les étapes suivantes :

10 a) la digestion des séquences d'ADN de mycobactéries par au moins une enzyme déterminée et la récupération des fragments de digestion obtenus ;

b) l'insertion des fragments de digestion dans un site de clonage compatible avec l'enzyme de l'étape a) du polylinker d'un vecteur selon l'une des revendications 1 à 13 ;

15 c) si besoin, l'amplification du fragment de digestion contenu dans le vecteur, par exemple par réplication de ce dernier après insertion du vecteur ainsi modifié dans une cellule déterminée, de préférence *E coli* ;

20 d) la transformation de cellules hôtes par le vecteur amplifié à l'étape c), ou en l'absence d'amplification, par le vecteur de l'étape b) ;

e) la culture de cellules hôtes transformées dans un milieu permettant la mise en évidence du marqueur d'exportation et/ou de sécrétion, et /ou du marqueur d'activité de promoteurs contenu dans le vecteur ;

25 f) la détection des cellules hôtes positives (colonies positives) pour l'expression du marqueur d'exportation et/ou de sécrétion, et /ou du marqueur d'activité de promoteurs ;

30 g) l'isolement de l'ADN des colonies positives et l'insertion de cet ADN dans une cellule identique à celle de l'étape c) ;

h) la sélection des insertions contenues dans le vecteur, permettant l'obtention de clones positifs pour le marqueur d'exportation et/ou de sécrétion, et /ou pour le marqueur d'activité de promoteurs ;

35 i) l'isolement et la caractérisation des fragments d'ADN de mycobactéries contenues dans ces insérats, et l'étape i) du procédé pouvant comporter en outre une étape de séquençage des

insertions sélectionnées.

16. Banque d'ADN génomique ou d'ADNc complémentaire d'ARNm de mycobactérie, caractérisée en ce qu'elle est obtenue par un  
5 procédé selon la revendication 14 et/ou un procédé comprenant les étapes a) et b), ou a), b) et c) du procédé selon la revendication 15.

17. Banque d'ADN génomique ou d'ADNc complémentaire d'ARNm de  
10 mycobactérie selon la revendication 16, caractérisée en ce que ladite mycobactérie est une mycobactérie pathogène.

18. Banque d'ADN génomique ou d'ADNc complémentaire d'ARNm de mycobactérie selon la revendication 17, caractérisée en ce que  
15 ladite mycobactérie est une mycobactérie appartenant au groupe du complexe *Mycobacterium tuberculosis*.

19. Banque d'ADN génomique ou d'ADNc complémentaire d'ARNm de mycobactérie selon la revendication 18, caractérisée en ce que  
20 ladite mycobactérie est *Mycobacterium tuberculosis*.

20. Séquence nucléotidique de mycobactérie susceptible d'être sélectionnée par un procédé selon l'une des revendications 14 et 15.

25

21. Séquence nucléotidique de mycobactérie selon la revendication 20 caractérisée en ce qu'elle est choisie parmi les séquences de fragments d'ADN de mycobactéries de séquence SEQ ID N°3, SEQ ID N°4, SEQ ID N°5, SEQ ID N°6, SEQ ID  
30 N°7, SEQ ID N°8, SEQ ID N°9 ou SEQ ID N°10, contenus respectivement dans les vecteurs p6D7 (CNCM, N°I-1814), p5A3 (CNCM, N°I-1815), p5F6 (CNCM, N°I-1816), p2A29 (CNCM, N°I-1817), p5B5 (CNCM, N°I-1819), p1C7 (CNCM, N°I-1820), p2D7 (CNCM, N°I-1821) et p1B7 (CNCM sous le N°I-1843).

35

22. Séquence nucléotidique de mycobactérie selon la revendication 20 caractérisée en ce qu'elle est choisie parmi les séquences de fragments d'ADN de mycobactéries de séquence nucléiques SEQ ID N°11 et SEQ ID N°24.
- 5
23. Polypeptide susceptibles d'être codés par une séquence nucléotidique de mycobactérie selon l'une des revendications 20 à 22.
- 10
24. Mycobactérie recombinantes caractérisée en ce qu'elle est transformée par un vecteur recombinant selon l'une des revendications 1 à 13.
25. Polynucléotide de séquence SEQ ID N°1 ou SEQ ID N°2.
- 15
26. Polynucléotide caractérisé en ce qu'il comprend un polynucléotide choisi parmi :
- a) un polynucléotide de séquence SEQ ID N°1 ou SEQ ID N°2,
- b) un polynucléotide dont la séquence est complémentaire de la
- 20 séquence du polynucléotide défini en a),
- c) un polynucléotide dont la séquence comporte au moins 80% d'homologie avec un polynucléotide défini en a) ou b),
- d) un polynucléotide hybridant dans des conditions de forte stringence avec une séquence de polynucléotide défini en a), b)
- 25 ou c),
- e) un fragment d'au moins 8 nucléotides consécutifs d'un polynucléotide défini en a), b), c) ou d).
27. Polynucléotide selon l'une des revendications 25 et 26,
- 30 caractérisé en ce que sa séquence nucléique est spécifique de séquence de mycobactéries appartenant au complexe de *Mycobacterium tuberculosis*.
28. Polypeptide de séquence d'acides aminés SEQ ID N°1 ou SEQ
- 35 ID N°2.

29. Polypeptide caractérisé en ce qu'il comprend un polypeptide choisi parmi :

a) un polypeptide de séquence d'acides aminés SEQ ID N°1 ou SEQ ID N°2,

5 b) un polypeptide homologue au polypeptide défini en a),

c) un fragment d'au moins 5 acides aminés d'un polypeptide défini en a) ou b),

d) un fragment biologiquement actif d'un polypeptide défini en a), b), ou c).

10

30. Polynucléotide caractérisé en ce qu'il code pour un polypeptide selon l'un des revendications 28 et 29.

15

31. Séquence d'acide nucléique utilisable comme amorce, caractérisée en ce que ladite séquence est choisie parmi les séquences d'acide nucléique de polynucléotide selon l'une des revendications 25 à 27, et 30.

20

32. Séquence d'acide nucléique selon la revendication 31, caractérisée en ce que ladite séquence est choisie parmi les séquences suivantes : SEQ ID N°25 et SEQ ID N°26.

25

33. Séquence d'acide nucléique selon l'une des revendications 31 et 32 pour la détection et/ou l'amplification de séquences nucléiques.

30

34. Séquence d'acide nucléique utilisable comme sonde, caractérisée en ce que ladite séquence est choisie parmi les séquences d'acide nucléique selon l'une des revendications 25 à 27, et 30.

35

35. Séquence d'acide nucléique selon la revendication 34, caractérisée en ce qu'elle est marquée par un composé radioactif ou par un composé non radioactif.

36. Séquence d'acide nucléique selon l'une des revendications 34 et 35, caractérisée en ce que celle-ci est immobilisée sur un support, de manière covalente ou non-covalente.
- 5 37. Séquence d'acide nucléique selon l'une des revendications 34 à 36 pour la détection et/ou l'amplification de séquences nucléiques.
38. Séquence d'acide nucléique selon l'une des revendications  
10 34 à 37, caractérisée en ce que ladite séquence est choisie parmi les séquences suivantes : du nucléotide 924 au nucléotide 1234 inclus dans SEQ ID N°1.
39. Vecteur recombinant de clonage, d'expression et/ou  
15 d'insertion, caractérisé en ce qu'il contient une séquence d'acide nucléique de polynucléotide selon l'une des revendications 25 à 27, et 30.
40. Vecteur recombinant selon la revendication 39, caractérisé  
20 en ce qu'il s'agit du plasmide pDP428 déposé le 28 janvier 1997 à la CNCM sous le N°I-1818.
41. Cellule hôte, caractérisée en ce qu'elle est transformée  
25 par un vecteur recombinant selon l'une des revendications 39 et 40.
42. Cellule hôte selon la revendication 41, caractérisée en ce qu'elle s'agit de la souche de E. coli transformée par le plasmide pDP428 déposé le 28 janvier 1997 à la CNCM sous le  
30 N°I-1818.
43. Polypeptide recombinant, caractérisé en ce qu'il est obtenu à partir d'une cellule hôte recombinant selon l'une des revendications 41 et 42.

44. Polypeptide hybride, caractérisé en ce qu'il comporte au moins la séquence d'un polypeptide selon l'une des revendications 28, 29 et 43 et une séquence d'un polypeptide susceptible d'induire une réponse immunitaire chez l'homme ou l'animal.

45. Polypeptide hybride selon la revendication 44, caractérisé en ce que le polypeptide susceptible d'induire une réponse immunitaire contient au moins un déterminant antigénique capable d'induire une réponse humorale et/ou cellulaire.

46. Polypeptide hybride selon l'une des revendications 44 et 45, caractérisé en ce que le déterminant antigénique correspond à un déterminant antigénique de protéines immunogènes sélectionnées par exemple parmi ESAT6 et DES de 45/47 kD de *Mycobacterium tuberculosis*.

47. Polypeptide hybride selon l'une des revendications 44 et 45, caractérisé en ce que le déterminant antigénique correspond à un déterminant antigénique de l'antigène de surface ou d'enveloppe d'un virus de l'hépatite ou de l'antigène VP1 du virus de la poliomyélite.

48. Polypeptide hybride selon l'une des revendications 44 et 45, caractérisé en ce que le déterminant antigénique est constitué de tout ou partie d'une protéine sélectionnée parmi les protéines gag, pol, nef ou env de HIV-1 ou HIV-2.

49. Polynucléotide codant pour un polypeptide hybride selon l'une des revendications 44 à 48.

50. Polypeptide hybride selon l'une des revendications 44 à 48 caractérisée en ce qu'il s'agit d'une protéine recombinante obtenue par l'expression d'un polynucléotide selon la revendication 49.

51. Procédé pour la détection *in vitro* d'anticorps dirigés contre une bactérie du complexe *Mycobacterium tuberculosis* dans



un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

- a) Mise en contact de l'échantillon biologique avec un polypeptide selon l'une des revendications 28, 29 et 43;
- 5      b) Mise en évidence du complexe antigène-anticorps formé.

52. Kit pour le diagnostic in vitro d'une infection par une mycobactérie appartenant au complexe *Mycobacterium tuberculosis*, comprenant:

- 10      a) un polypeptide selon l'une des revendications 28, 29 et 43;
- b) les réactifs pour la constitution du milieu propice à la réaction immunologique ;
- c) les réactifs permettant la détection des complexes
- 15      antigène-anticorps produits par la réaction immunologique ;
- d) le cas échéant, un échantillon biologique de référence (témoin négatif) dépourvu d'anticorps reconnus par ledit polypeptide ;
- e) le cas échéant, un échantillon biologique de
- 20      référence (témoin positif) contenant une quantité prédéterminée d'anticorps reconnus par ledit polypeptide.

53. Anticorps mono- ou polyclonaux, leurs fragments, ou anticorps chimériques, caractérisés en ce qu'ils sont capables

25      de reconnaître spécifiquement un polypeptide selon l'une des revendications 28, 29 et 43.

54. Anticorps selon la revendication 53, caractérisé en ce qu'il s'agit d'un anticorps marqué.

30

55. Procédé pour la détection spécifique de la présence d'un antigène d'une bactérie du complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

- 35      a) Mise en contact de l'échantillon biologique avec un anticorps selon l'une des revendications 53 et 54;
- b) Mise en évidence du complexe antigène-anticorps formé.

56. Kit pour la détection spécifique de la présence d'un antigène d'une bactérie du complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants :

- 5       a) Un anticorps polyclonal ou monoclonal selon l'une des revendications 53 et 54;
- b) les réactifs pour la constitution du milieu propice à la réaction immunologique ;
- c) les réactifs permettant la détection des complexes  
10 antigène-anticorps produits par la réaction immunologique.

57. Procédé de détection et d'identification rapide de *M. tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comporte les étapes suivantes :

- 15       a) Isolement de l'ADN à partir de l'échantillon biologique à analyser, ou obtention d'un ADNc à partir de l'ARN de l'échantillon biologique ;
- b) Amplification spécifique de l'ADN des mycobactéries appartenant au complexe *Mycobacterium tuberculosis* à l'aide  
20 d'amorces selon l'une des revendications 31 à 33;
- c) Analyse des produits d'amplification.

58. Procédé pour la détection spécifique de bactéries appartenant au complexe *Mycobacterium tuberculosis* dans un  
25 échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

- a) Mise en contact d'une sonde oligonucléotidique selon l'une des revendications 34 à 38 avec un échantillon biologique, l'ADN contenu dans l'échantillon biologique ayant,  
30 le cas échéant, préalablement été rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de la sonde à l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*;
- b) Détection de l'hybride formé entre la sonde  
35 oligonucléotidique et l'ADN de l'échantillon biologique.

59. Procédé pour la détection spécifique de bactéries appartenant au complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

5     a) Mise en contact d'une sonde oligonucléotidique immobilisée sur un support selon la revendication 36 avec un échantillon biologique, l'ADN de l'échantillon, ayant, le cas échéant, été préalablement rendu accessible à l'hybridation, dans des conditions permettant l'hybridation de la sonde à  
10 l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*;

      b) Mise en contact de l'hybride formé entre la sonde oligonucléotidique immobilisée sur un support et l'ADN contenu dans l'échantillon biologique, le cas échéant après élimination de l'ADN de l'échantillon biologique n'ayant pas  
15 hybridé avec la sonde, avec une sonde oligonucléotidique marquée selon la revendication 35.

60. Procédé de détection selon la revendication 59, caractérisé en ce que, préalablement à l'étape a), l'ADN de l'échantillon  
20 biologique, ou l'ADNc obtenu par transcription inverse de l'ARN de l'échantillon, est amplifié à l'aide d'un couple d'amorces selon l'une des revendications 31 à 33.

61. Procédé pour la détection spécifique de la présence d'une  
25 bactérie appartenant au complexe de *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

      a) Mise en contact de l'échantillon biologique avec un couple d'amorces selon l'une des revendications 31 à 33, l'ADN  
30 contenu dans l'échantillon ayant été, le cas échéant, préalablement rendu accessible à l'hybridation, dans des conditions permettant une hybridation des amorces à l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*;

      b) amplification de l'ADN de la bactérie du complexe  
35 *Mycobacterium tuberculosis*;

c) mise en évidence de l'amplification de fragments d'ADN correspondant au fragment encadré par les amorces, par exemple par électrophorèse sur gel ou au moyen d'une sonde oligonucléotidique marquée selon la revendication 35.

5

62. Procédé pour la détection spécifique de la présence d'une bactérie appartenant au complexe de *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les étapes suivantes :

10 a) Mise en contact de l'échantillon biologique avec deux couples d'amorces selon l'une des revendications 31 à 33, l'ADN contenu dans l'échantillon ayant été, le cas échéant, préalablement rendu accessible à l'hybridation, dans des conditions permettant une hybridation des amorces à l'ADN  
15 d'une bactérie du complexe *Mycobacterium tuberculosis*;

b) amplification de l'ADN de la bactérie du complexe *Mycobacterium tuberculosis*;

c) mise en évidence de l'amplification de fragments d'ADN correspondant au fragment encadré par lesdites amorces, par  
20 exemple par électrophorèse sur gel ou au moyen d'une sonde oligonucléotidique marquée selon la revendication 35.

63. Kit pour la détection de la présence d'une bactérie du complexe *Mycobacterium tuberculosis* dans un échantillon  
25 biologique, caractérisé en ce qu'il comprend les éléments suivants :

a) Une sonde oligonucléotidique selon l'une des revendications 34 à 38;

b) Les réactifs nécessaires à la mise en oeuvre d'une  
30 réaction d'hybridation;

c) Le cas échéant, un couple d'amorces selon l'une des revendications 31 à 33 ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN (ADN génomique, plasmidique ou ADNc) d'une bactérie du complexe *Mycobacterium*  
35 *tuberculosis*.

64. Kit ou nécessaire pour la détection de la présence d'une bactérie du complexe *Mycobacterium tuberculosis* dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants :

- 5       a) Une sonde oligonucléotidique, dite sonde de capture, selon la revendication 36;
- b) Une sonde oligonucléotidique, dite sonde de révélation, selon l'une des revendications 34 à 38.
- c) Le cas échéant, un couple d'amorces selon l'une des  
10       revendications 31 à 33 ainsi que les réactifs nécessaires à une réaction d'amplification de l'ADN d'une bactérie du complexe *Mycobacterium tuberculosis*.

65. Kit pour l'amplification de l'ADN d'une bactérie du  
15       complexe *Mycobacterium tuberculosis* présent dans un échantillon biologique, caractérisé en ce qu'il comprend les éléments suivants :

- a) Un couple de fragments d'acide nucléique selon l'une des revendications 31 à 33;
- 20       b) Les réactifs nécessaires pour effectuer une réaction d'amplification d'ADN;
- c) Eventuellement un composant permettant de vérifier la séquence du fragment amplifié, plus particulièrement une sonde oligonucléotidique selon l'une des revendications 34 à 38.

25

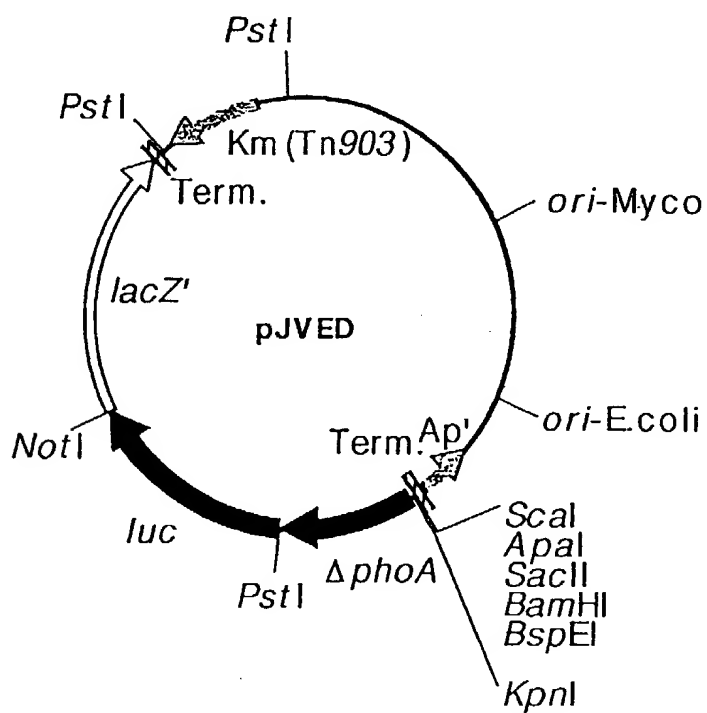
66. Composition immunogène caractérisée en ce qu'elle comprend un ou plusieurs polypeptides selon l'une des revendications 28, 29 et 43 et/ou un ou plusieurs polypeptides hybrides selon l'une des revendications 44 à 50.

30

67. Vaccin caractérisé en ce qu'il contient un ou plusieurs polypeptides selon l'une des revendications 28, 29 et 43 et/ou un ou plusieurs polypeptides hybrides selon l'une des revendications 44 à 50, en association avec un véhicule  
35       pharmaceutiquement compatible et, le cas échéant un adjuvant de l'immunité approprié.

68. Vaccin destiné à l'immunisation à l'encontre d'une infection bactérienne ou virale, telle que la tuberculose ou l'hépatite, comprenant un vecteur selon la revendication 39 ou un polynucléotide selon la revendication 49, en association  
5 avec un véhicule pharmaceutiquement compatible.

A



B

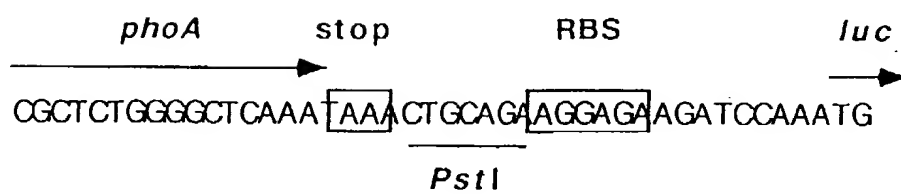


FIGURE 1

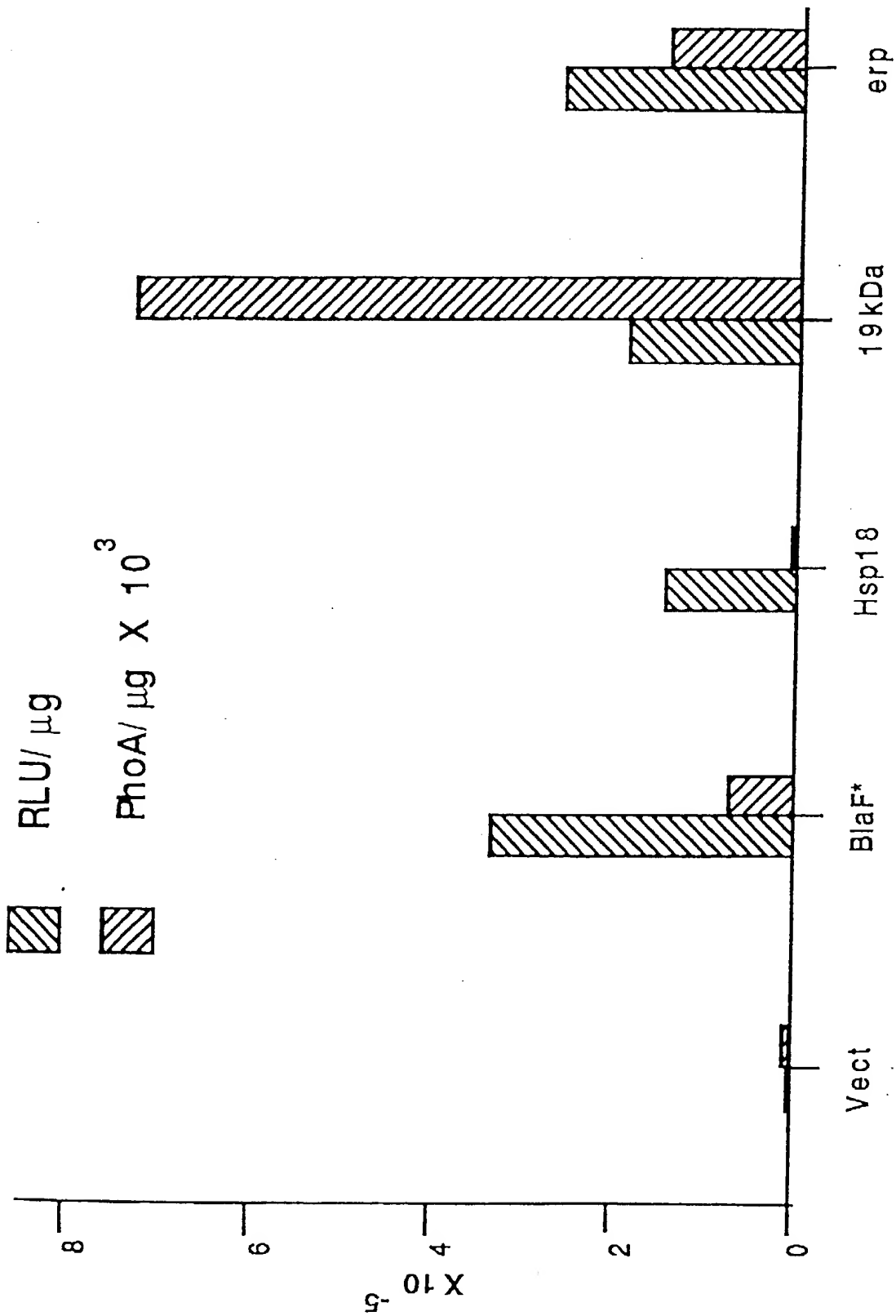
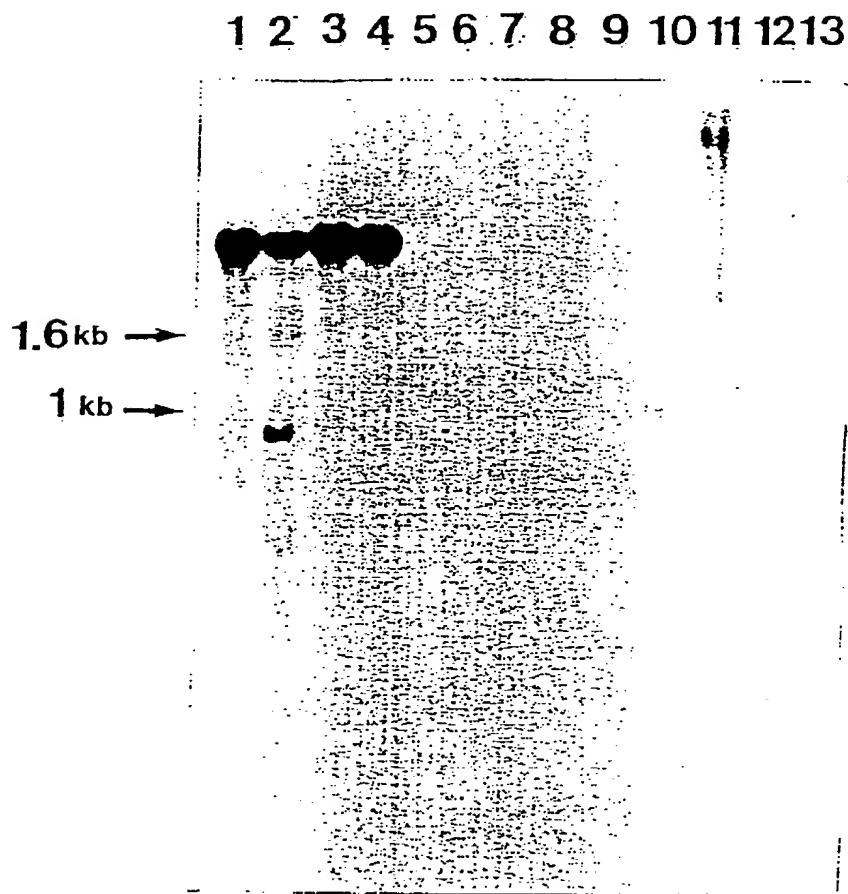


FIGURE 2





1: *M. tuberculosis* 2: *M. bovis* 3: BCG 4: *M. africanum* 5: cancelled 6: *M. fortuitum* 7: *M. simiae* 8: *M. avium* 9: *M. chelonae* 10: *M. flavescens* 11: *M. gordonae* 12: *M. marinum* 13: *M. kansasii*

FIGURE 3

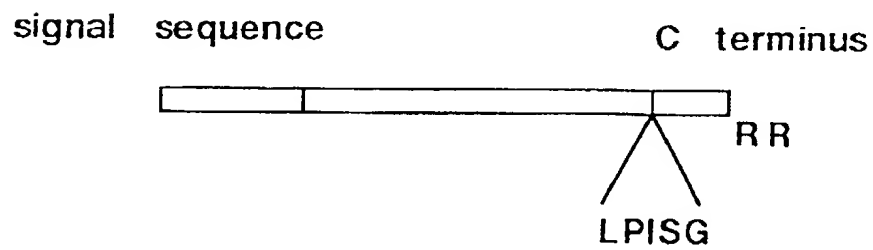
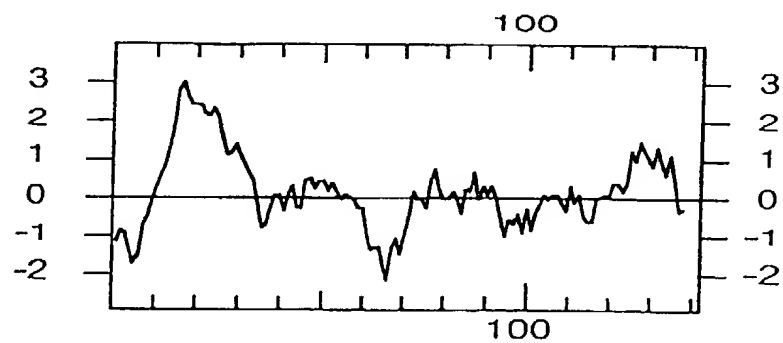


FIGURE 4

ins.428def.strider -&gt; Genes

DNA sequence 1243 b.p. GGATCCAGGGA ... CGCATCGGATCC linear

```

1  GGATCCAGGGAACGTGACC ATG GTC GTA GGG ATG ACT TGA CAGTTTCAACGGGGTGGACCAACCGTTGCGC 72
1  M V V G M T * 7
73 TCAGAAGGCATACGTTGGTGAACACGTCGGAAAGCTGGGAGGTGAATCTG ATG GCT GGC GAC CAA GAG CTG 144
1  M A G D Q E L 7
145 GAA CTG CCG TTC GAC GTT CCT CTT TAC ACG CTT GCC GAG GCA TCG CCG TAC CTG GTG GTT 204
8  E L R F D V P L Y T L A E A S R Y L V V 27
205 CCC CCG GCC ACC CTG GCT ACG TGG GCT GAC GGC TAC GAG CGT CCG CCG GCC AAC GCA CCG 264
28  P R A T L A T W A D G Y E R R P A N A P 47
265 GCG GTC CAG GGG CAA CCG ATC GCC TTT GAC GGC TAT TCG GTC GCG CAG CTT TTT GGC GAC 324
48  A V Q G Q P I A F D A Y S V A Q L F G D 67
325 GTC ACT GGT GCC CCG GTT GCG GGC GTC CAG CCG CAG CGA CAC CAC ATA CCG CCG GTC CCG 384
68  V T G A R V A G V Q P Q R H H I R P V R 87
385 TTG CCG GGG CCG TTG GGT GGG GTT GGG TGC CTC CGT CAC CCG AGG CAG TTC GCT GGC TAT 444
88  L R G P L G G V G C L R H P R Q F A G Y 107
445 TTG TCG CAG TAG CCGCAGCGCATTTGTCG ATG TCT TGG TAG CTAGCATCGGTGGGGGGCGCTACAGCG 515
108  L S Q * M S W * 4
516 CCAGCGCGGGGCTCCCGGTCGGGTAGTGGCGGTGAGTTGGTGGTGGACAGCA ATG ACT GCG ACC CCG 587
1  M T A T R 5
588 CGA CTT CGA AAC CCG CAC CCG TTA GAT TCC CCG ACT GCG TCA TCG CCA GGT AAA CCG CCG 647
6  R L R N R H R L D S P T A S S P G K P P 25
648 GCA CTA ACG CCA GCA ACC AAC CCG TGA AGACCAACCAACCGGCACCTGGCAGGTTGGCGGCTCAACCGCATC 718
26  A L T P A T N P * 34
719 ATG AAC TGC TGG AIT TCG GAC TCC CCG TAC TCT CCG GCA GTG CGT GCC CCG GAG CCT ACC 778
1  M N C W I S D S P Y S R A V R A R E P T 20
779 GAA GAT CCG GTG CAT GCG TTC GGC GTG GAC CCG ACA GCA CCT GGA GTT GGC GGC GCC GAG 838
21  E D R V H A F G V D R T A P G V G G A E 40
839 GGC CGA GAT GGC AGG ATG ACG GAT CGT CCG GGG CCG GAA CTC CCA GGC CCG CCG ACC GTC 898
41  G R D G R M T D R R G R E L P G R R T V 60
899 GCA AAC CCG TCG CAA ACC CGT CCG AAA CCG TAA GGAGTCATCC ATG AAG ACA GGC ACC GCG 959
61  A N P S Q T R R K P * M K T G T A 6
960 ACC ACG CCG CCG AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CCG GGG GCC GCC GTT 1019
7  T T R R R L L A V L I A L A L P G A A V 26
1020 GCG CTG CTG GCC GAA CCA TCA GCG ACC GGC GCG TCG GAC CCG TCG GCG GCC AGC GAA GTG 1079
27  A L L A E P S A T G A S D P C A A S E V 46
1080 GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT TCA CAC CCA GAG 1139
47  A R T V G S V A K S M G D Y L D S H P E 66
1140 ACC AAC CAG GTG ATG ACC GCG GTC TTG CAG CAG CAG GTA GGG CCG GGG TCG GTC CCA TCG 1199
67  T N Q V M T A V L Q Q Q V G P G S V A S 86
1200 CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT CC 1243
87  L K A H F E A N P K V A S D 100

```

SEQ ID N° 1

FIGURE 5

491 CCGGTGGGGGGGCGCTACCAGCGCCAGCGCGGGGCTCCCGGTCCGGGTA GTG CGC GTC GAG TTG GTC GTG 563  
 1 V R V E L V V 7  
 564 GAC CAG CAA TGA CTGCGACCCGGGACTTGGAAACCGCCACCGGTTAGATTCCCGACTGCGTCATCCAGGTAA 639  
 8 D Q Q \*  
 640 ACCGCGGGCACTAACGCCAGCAACCAACCC GTG AAG ACC AAC CAA CGG CAC CTG CGC AGG TTG CGG 705  
 1 V R T N Q R H L R R L R 12  
 706 CTC AAC CGC ATC ATG AAC TGC TGG ATT TCG GAC TCC CCG TAC TCT CGC GCA GTG CGT GCC 765  
 13 L N R I H N C W I S D S P Y S R A V R A 32  
 766 CGC GAG CCT ACC GAA GAT CGC GTG CAT GCG TTC GGC GTG GAC CGC ACA GCA CCT GGA GTT 825  
 33 R E P T E D R V H A F G V D R T A P G V 52  
 826 GGC GGC GGC GAG GGC CGA GAT GGC AGG ATG ACG GAT CGT CGG GGG CGG GAA CTC OCA GGC 885  
 53 G G A E G R D G R M T D R R G R E L P G 72  
 886 CGC CGG ACC GTC GCA AAC CGG TCG CAA ACC CGT CGC AAA CGG TAA GGAGTCATCC ATG AAG 946  
 73 R R T V A N P S Q T R R K P \* XXXXXX M R 2  
 947 ACA GGC ACC GCG ACG ACG CGG CGC AGG CTG TTG GCA GTA CTG ATC GCC CTC GCG TTG CGG 1006  
 3 T G T A T T R R R L L A V L I A L A L P 22  
 1007 GGC GCC GGC GTT GCG CTG CTG GGC GAA CCA TCA GCG ACC GGC GCG TCG GAC CGG TGC GCG 1066  
 23 G A A V A L L A E P S A T G A S D P C A 42  
 1067 GCC AGC GAA GTG GCG AGG ACG GTC GGT TCG GTC GCC AAG TCG ATG GGC GAC TAC CTG GAT 1126  
 43 A S E V A R T V G S V A K S M G D Y L D 62  
 1127 TCA CAC CCA GAG ACC AAC CAG GTG ATG ACC CCG GTC TTG CAG CAG CAG GTA GGG CCG GGG 1186  
 63 S H P E T N Q V H T A V L Q Q Q V G P G 82  
 1187 TCG GTC GCA TCG CTG AAG GCC CAT TTC GAG GCG AAT CCC AAG GTC GCA TCG GAT CTG CAC 1246  
 63 S V A S L K A H F E A N F K V A S D L H 102  
 1247 GCG CTT TCG CAA CCG CTG ACC GAT CTT TCG ACT CCG TGC TCG CTG CCG ATC AGC GGC CTG 1306  
 103 A L S Q P L T D L S T R C S L P I S G L 122  
 1307 CAG GCG ATC GGT TTG ATG CAG GCG GTG CAG GGC GCC CGC CGG TAG ATG CCG GAC CGC CGC 1366  
 123 Q A I G L M O A V Q G A R R I \* H P D R R 5  
 1367 CCG GTC CCG CGC AGT CGA CGT GAG GCA GCG GTC GCC TAC CCG GGC GGT GTC TCG CCG CCT 1426  
 6 R V R R S R R E A A V A Y R G G V S P P 25  
 1427 TCT GGT CGC AGG TCA GGG GTC GGC GCT GGA CCT TGC GGT GTG GTT TCG ACC GGC TCG TCG 1486  
 26 S G R R S G V G A G P C G V V S T G S S 45  
 1487 CAG GGT GTG CCC TGC GGT TGG ATG ACA AGT CGC AGG TTT GGA TCG GTT GGC GGG TCG CGA 1546  
 46 Q G V P C G W M T S R R F G S V G G S R 65  
 1547 TCG TTG T 1553  
 66 S L 67

SEQ ID N° 2

FIGURE 6

## 6D7.95-Seq.11 -&gt; 1-phase Translation

DNA sequence 304 b.p. TCGCCGGCTCCG ... CGCATGCCGATC linear

```

1/1                                31/11
TCG CCG GCT CGC GGA CGT AGA TAA TAG CTC ACC GTT GGA CGA CCT CGA CAG GGT CCT TTG
ser pro ala arg gly arg arg OCH AMB leu thr val gly arg pro arg gln gly pro leu
61/21                                91/31
TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG TCG GGT CAT CGC CTA AGG CTA CCG TTC
OPA leu pro gly leu thr arg thr thr thr glu ser gly his arg leu arg leu pro phe
121/41                                151/51
TGA CCT GGG GTG CGT GGG CGC CGA CGA GTG AGG CAG TCA TGT CTC AGG GCC CAC CGC CAC
OPA pro gly val arg gly arg arg arg val arg gln ser cys leu arg ala his arg his
181/61                                211/71
CTC GGT CGC CGG CAG TGT CAG CAT GTG CAG ATG ACT CCA CGC AGC TTG TTC GTG TTG GTG
leu gly arg arg gln cys gln his val gln met thr pro arg ser leu phe val leu val
241/81                                271/91
TCG TGG TTG CGA CGA CTT GGC GCT GGT GAG CGC ACC CGC CGG CGT CGT GCC CGC CAT GCG
ser trp leu arg arg leu gly ala gly glu arg thr arg arg arg arg ala ala his ala
301/101
GAT C
asp

```

SEQ ID N° 3

FIGURE 7A

## 6D7.95-Seq.11 [2 to 304] -&gt; 1-phase Translation

DNA sequence 304 b.p. TCGCCGGCTCGC ... CGCATGCCGATC linear

```

2/1                               32/11
CGC CGG CTC GCG GAC GTA GAT AAT AGC TCA CCG TTG GAC GAC CTC GAC AGG GTC CTT TGT
arg arg leu ala asp val asp asn ser ser pro leu asp asp leu asp arg val leu cys
62/21                             92/31
GAC TGC CCG GCT TGA CGC GGA CGA CCA CAG AGT CCG GTC ATC GCC TAA GGC TAC CGT TCT
asp cys arg ala OPA arg gly arg pro gln ser arg val ile ala OCH gly tyr arg ser
122/41                           152/51
GAC CTG GGG TGC GTG GGC GCC GAC GAG TGA GGC AGT CAT GTC TCA GGC CCC ACC GCC ACC
asp leu gly cys val gly ala asp glu OPA gly ser his val ser gly pro thr ala thr
182/61                           212/71
TCG GTC GCC GGC AGT GTC AGC ATG TGC AGA TGA CTC CAC GCA GCT TGT TCG TGT TCG TGT
ser val ala gly ser val ser met cys arg OPA leu his ala ala cys ser cys trp cys
242/81                           272/91
CGT GGT TGC GAC GAC TTG GCG CTG GTG AGC GCA CCC GCC GGC GTC GTG CCG CGC ATG CGG
arg gly cys asp asp leu ala leu val ser ala pro ala gly val val pro arg met arg
302/101
ATC
ile

```

SEQ ID N° 3

FIGURE 7B

## 6D7.95-Seq.11 [3 to 304] -&gt; 1-phase Translation

DNA sequence 304 b.p. TCGCCGCTCGC ... CGCATGCCGATC linear

3/1 33/11  
 GCC GGC TCG CGG ACG TAG ATA ATA GCT CAC CGT TGG ACG ACC TCG ACA GGG TCC TTT GTG  
 ala gly ser arg thr AMB ile ile ala his arg trp thr thr ser thr gly ser phe val  
 63/21 93/31  
 ACT GGC GGG CTT GAC GCG GAC GAC CAC AGA GTC GGG TCA TCG CCT AAG GCT ACC GTT CTG  
 thr ala gly leu asp ala asp asp his arg val gly ser ser pro lys ala thr val leu  
 123/41 153/51  
 ACC TGG GGT GCG TCG GCG CCG ACG AGT GAG GCA GTC ATG TCT CAG GGC CCA CCG CCA CCT  
 thr trp gly ala trp ala pro thr ser glu ala val met ser gln gly pro pro pro pro  
 183/61 213/71  
 CCG TCG CCG GCA GTG TCA GCA TGT GCA GAT GAC TCC ACG CAG CTT GTT CGT GTT GGT GTC  
 arg ser pro ala val ser ala cys ala asp asp ser thr gln leu val arg val gly val  
 243/81 273/91  
 GTG GTT GCG ACG ACT TGG CCG TGG TGA GCG CAC CCG CCG GCG TCG TGC CGC GCA TGC GGA  
 val val ala thr thr trp arg trp OPA ala his pro pro ala ser cys arg ala cys gly

TC

SEQ ID N° 3

FIGURE 7C

## 5A3.95-Seq.16 -&gt; 1-phase Translation

DNA sequence 389 b.p. CCAATTTTCTT ... ACGACCACGATC linear

```

1/1                                31/11
CCA ATT TTC CTT CGC GCC GTG CAA TAC CAT CTG CAA GAC CAG CGA CGG CCC GTG GTT GCG
pro ile phe leu arg ala val gln tyr his leu gln asp gln arg arg pro val val ala
61/21                                91/31
GTC GCG CAG CTT GCG GAA ACC GGG TAT GGA CCC TGC CGT ACC GTT GTT GCC ACT TGA TGT
val ala gln leu ala glu thr gly tyr gly pro cys arg thr val val ala thr OPA cys
121/41                               151/51
CGT CGC TCT CCA CCC GTC GGG GGG CGA AAG CCA TTC CGA CAC TGG GAT CCT CAA AAC GTC
arg arg ser pro pro val gly gly arg lys pro phe arg his trp asp pro gln asn val
181/61                               211/71
GGC TGA GTG TCT GCA GGG CTC CGG GGA GCA GCC GAT CAT CAC CAT GTA CGA ACT GAA TAA
gly OPA val ser ala gly leu arg gly ala ala asp his his his val arg thr glu OCH
241/81                               271/91
GTC CCC CGC GCG CGA CTT CCA GAC ATT TGT TGT GGT TTC GGT TGA GGC CGA GGC GAG GCT
val pro arg ala arg leu pro asp ile cys cys gly phe gly OPA gly arg gly glu ala
301/101                              331/111
CAT TTC GCA GCA ACC GGT CTC CGG GTC GCA GCA TCG TTG CGG CGA TCG CGG CGC AGT CGT
his phe ala ala thr gly leu arg val ala ala ser leu arg arg ser arg arg ser arg
361/121
CGG ACG AGT CGT CGT CAA CGA CCA CGA TC
arg thr ser arg arg gln arg pro arg

```

SEQ ID N° 4

FIGURE 8A



## 5A3.95-Seq.16 [2 to 389] -&gt; 1-phase Translation

DNA sequence 389 b.p. CCAATTTCCTT ... ACGACCACGATC linear

2/1  
 CAA TTT TCC TTC GCG CCG TGC AAT ACC ATC TGC AAG ACC AGC GAC GGC CCG TGG TTG CCG  
 gln phe ser phe ala pro cys asn thr ile cys lys thr ser asp gly pro trp leu arg  
 62/21  
 TCG CGC AGC TTG CCG AAA CCG GGT ATG GAC OCT GCC GTA CCG TTG TTG CCA CTT GAT GTC  
 ser arg ser leu arg lys pro gly met asp pro ala val pro leu leu pro leu asp val  
 122/41  
 GTC GCT CTC CAC CCG TCG GGG GGC GAA AGC CAT TCC GAC ACT GGG ATC CTC AAA ACG TCG  
 val ala leu his pro ser gly gly glu ser his ser asp thr gly ile leu lys thr ser  
 182/61  
 GCT GAG TGT CTG CAG GGC TCC GGG GAG CAG CCG ATC ATC ACC ATG TAC GAA CTG AAT AAG  
 ala glu cys leu gln gly ser gly glu gln pro ile ile thr met tyr glu leu asn lys  
 242/81  
 TCC CCC GCG CCG GAC TTC CAG ACA TTT GTT GTG GTT TCG GTT GAG GCC GAG GCG AGG CTC  
 ser pro ala arg asp phe gln thr phe val val val ser val glu ala glu ala arg leu  
 302/101  
 ATT TCG CAG CAA CCG GTC TCC GGG TCG CAG CAT CCG TGC GGC GAT CCG GGC GCA GTC GTC  
 ile ser gln gln pro val ser gly ser gln his arg cys gly asp arg gly ala val val  
 362/121  
 GGA CGA GTC GTC GTC AAC GAC CAC GAT C  
 gly arg val val val asn asp his asp

SEQ ID N° 4

FIGURE 8B

## 5A3.95-Seq.16 [3 to 389] -&gt; 1-phase Translation

DNA sequence 389 b.p. CCAATTTTCCTT ... ACGACCACGATC linear

```

3/1
AAT TTT OCT TCG CGC CGT GCA ATA CCA TCT GCA AGA CCA GCG ACG GCC CGT GGT TGC GGT
asn phe pro ser arg arg ala ile pro ser ala arg pro ala thr ala arg gly cys gly
63/21
CGC GCA GCT TGC GGA AAC CGG GTA TGG ACC CTG CCG TAC CGT TGT TGC CAC TTG ATG TCG
arg ala ala cys gly asn arg val trp thr leu pro tyr arg cys cys his leu met ser
123/41
TCG CTC TCC ACC CGT CGG GGG GCG AAA GCC ATT CCG ACA CTG GGA TCC TCA AAA CGT CGG
ser leu ser thr arg arg gly ala lys ala ile pro thr leu gly ser ser lys arg arg
183/61
CTG AGT GTC TGC AGG GCT CCG GGG AGC AGC CGA TCA TCA CCA TGT ACC AAC TGA ATA AGT
leu ser val cys arg ala pro gly ser ser arg ser ser pro cys thr asn opa ile ser
243/81
CCC CCG CGC GCG ACT TCC AGA CAT TTG TTG TGG TTT CCG TTG AGG CCG AGG CGA GGC TCA
pro pro arg ala thr ser arg his leu leu trp phe arg leu arg pro arg arg gly ser
303/101
TTT CCG AGC AAC CGG TCT CCG GGT CCG AGC ATC GTT CCG GCG ATC GCG GCG CAG TCG TCG
phe arg ser asn arg ser pro gly arg ser ile val ala ala ile ala ala gln ser ser
363/121
GAC GAG TCG TCG TCA ACG ACC ACG ATC
asp glu ser ser ser thr thr thr ile

```

SEQ ID N° 4

FIGURE 8C

## 5F6 -&gt; 1-phase Translation

DNA sequence 316 b.p. GATCGCGGTCAA ... CCGTCGCGGATC linear

1/1  
 GAT CGC GGT CAA CGA GGC CGA ATA CGG CGA GAT GTG GGC CCA AGA CGC CGC CGC GAT GTT  
 asp arg gly gln arg gly arg ile arg arg asp val gly pro arg arg arg arg asp val  
 91/31  
 61/21  
 TGG CTA CGC CGC GGC GAC GGC GAC GGC GAC GGC GAC GTT GCT GCC GTT CGA GGA GGC GCC  
 trp leu arg arg gly asp gly asp gly asp gly asp val ala ala val arg gly gly ala  
 151/51  
 121/41  
 GGA GAT GAC CAG CGC GGG TGG GCT CCT CGA GCA GGC CGC CGC GGT CGA GGA GGC CTC CGA  
 gly asp asp gln arg gly trp ala pro arg ala gly arg arg gly arg gly gly leu arg  
 211/71  
 181/61  
 CAC CGC CGC GGC GAA CCA GTT GAT GAA CAA TGT GCC CCA GGC GCT CCA ACA GCT GGC CCA  
 his arg arg gly glu pro val asp glu gln cys ala pro gly ala ala thr ala gly pro  
 271/91  
 241/81  
 GCC CAC GCA GGG CAC CAC GCC TTC TTC CAA GCT GGG TGG CCT GTG GAA GAC GGT CTC GCC  
 ala his ala gly his his ala phe phe gln ala gly trp pro val glu asp gly leu ala  
 301/101  
 GCA TCG GTC GCC GAT C  
 ala ser val ala asp

SEQ ID N° 5

FIGURE 9A

## 5F6 [2 to 316] -&gt; 1-phase Translation

DNA sequence 316 b.p. GATCGCGGTCAA ... CCGTCGCCGATC linear

```

2/1                               32/11
ATC GCG GTC AAC GAG GCC GAA TAC GGC GAG ATG TGG GCC CAA GAC GCC GCC GCG ATG TTT
ile ala val asn glu ala glu tyr gly glu met trp ala gln asp ala ala ala met phe
62/21                             92/31
GGC TAC GCC GCG GCG ACG GCG ACG GCG ACG GCG ACG TTG CTG CCG TTC GAG GAG GCG CCG
gly tyr ala ala ala thr ala thr ala thr ala thr leu leu pro phe glu glu ala pro
122/41                           152/51
GAG ATG ACC AGC GCG GGT GGG CTC CTC GAG CAG GCC GCC GCG GTC GAG GAG GCC TCC GAC
glu met thr ser ala gly gly leu leu glu gln ala ala ala val glu glu ala ser asp
182/61                           212/71
ACC GCC GCG GCG AAC CAG TTG ATG AAC AAT GTG CCC CAG GCG CTG CAA CAG CTG GCC CAG
thr ala ala ala asn gln leu met asn asn val pro gln ala leu gln gln leu ala gln
242/81                           272/91
CCC ACG CAG GGC ACC ACG CCT TCT TCC AAG CTG GGT GGC CTG TGG AAG ACG GTC TCG CCG
pro thr gln gly thr thr pro ser ser lys leu gly gly leu trp lys thr val ser pro
302/101
CAT CCG TCG CCG ATC
his arg ser pro ile

```

SEQ ID N° 5

FIGURE 9B

## 5F6 [3 to 316] -&gt; 1-phase Translation

DNA sequence 316 b.p. GATCGCGGTCAA ... CGGTCGCCGATC linear

3/1 33/11  
 TCG CGG TCA ACG AGG CCG AAT ACG GCG AGA TGT GGG CCC AAG ACG CCG CCG CGA TGT TTG  
 ser arg ser thr arg pro asn thr ala arg cys gly pro lys thr pro pro arg cys leu  
 63/21 93/31  
 GCT ACG CCG CCG CGA CCG CGA CCG CGA CCG CGA CGT TGC TGC CGT TCG AGG AGG CGC CGG  
 ala thr pro arg arg arg arg arg arg arg arg arg cys cys arg ser arg arg arg arg  
 123/41 153/51  
 AGA TGA CCA GCG CCG GTG GGC TCC TCG AGC AGG CCG CCG CCG TCG AGG AGG CCT CCG ACA  
 arg OPA pro ala arg val gly ser ser ser arg pro pro arg ser arg arg pro pro thr  
 183/61 213/71  
 CCG CCG CCG CGA ACC AGT TGA TGA ACA ATG TGC CCC AGG CGC TGC AAC AGC TGG CCC AGC  
 pro pro arg arg thr ser OPA OPA thr met cys pro arg arg cys asn ser trp pro ser  
 243/81 273/91  
 CCA CGC AGG GCA CCA CGC CTT CTT CCA AGC TGG GTG GCC TGT GGA AGA CGG TCT CGC CGC  
 pro arg arg ala pro arg leu leu pro ser trp val ala cys gly arg arg ser arg arg  
 303/101  
 ATC GGT CGC CGA TC  
 ile gly arg arg

SEQ ID N° 5

FIGURE 9C

## -2A29.96-Seq.17 -&gt; 1-phase Translation

DNA sequence 433 b.p. GGATCCTGATGC ... AAGGAGAAGATC linear

```

1/1                                31/11
GGA TCC TGA TGC AAG TGG TCC GGG ATT TGT CCG CAG CCA CGG CGG TCC CGT CGA CCA ACG
gly ser OPA cys lys trp ser gly ile cys arg gln pro arg arg ser arg arg pro thr
61/21                                91/31
TTG GTG CAT CCG GGC TGC GAG CAT GCA CGC ACC GAC CAG CGC GGC GAG CGC GGC TAG CTG
leu val his pro gly cys glu his ala arg thr asp gln arg gly glu arg gly AMB leu
121/41                                151/51
CTT GCC CAC TGT TCC TCC CTG CCG GCA CCA TGT CCG ACA AGC TTA AGC GCA GCA GTA CCG
leu ala his cys ser ser leu pro ala pro cys ala thr ser leu ser ala ala val pro
181/61                                211/71
GCG GTG CCT GGG CAT CCA GCA AAA CGG GGA GCT CAA GAA CGA TTC ATG AAC GAG GCG TCG
ala val pro gly his pro ala lys arg gly ala gln glu arg phe met asn glu gly ser
241/81                                271/91
TCA CCA ACG TCG AAA CCG ACG GTT GCC AGC CCG CCC ACG ATA TTG CGT GCT CGA GCG TCC
ser pro thr ser lys pro thr val ala ser arg pro thr ile leu arg ala arg gly ser
301/101                               331/111
GCT GTA CCC TCA CCG AAC GTG AGT CCC ACA CCG CCG AGG CCG CCG ACT CTG CCG TCG TTA
ala val pro ser pro asn val ser pro thr pro arg arg arg ala thr leu ala ser leu
361/121                               391/131
GCA GCC GAG CTC AAG GTG TCC CGC ACC ACT GTC TCG AAT GCT TTT AAC CGA CCG GAT CCA
ala ala glu leu lys val ser arg thr thr val ser asn ala phe asn arg pro asp pro
421/141
GAA GGA GAA GAT C
glu gly glu asp

```

SEQ ID N° 6

FIGURE 10A

## -2A29.96-Seq.17 [2 to 433] -&gt; 1-phase Translation

DNA sequence 433 b.p. GGATCCTGATGC ... AAGGAGAAGATC linear

```

2/1                                32/11
GAT CCT GAT GCA AGT GGT CCG GGA TTT GTC GGC AGC CAC GGC GGT CCC GTC GAC CAA CGT
asp pro asp ala ser gly pro gly phe val gly ser his gly gly pro val asp gln arg
62/21                                92/31
TGG TGC ATC CCG GCT GCG AGC ATG CAC GCA CCG ACC AGC GCG GCG AGC GCG GCT AGC TGC
trp cys ile arg ala ala ser met his ala pro thr ser ala ala ser ala ala ser cys
122/41                                152/51
TTG CCC ACT GTT CCT CCC TGC CCG CAC CAT GTG CGA CAA GCT TAA GCG CAG CAG TAC CGG
leu pro thr val pro pro cys arg his his val arg gln ala OCH ala gln gln tyr arg
182/61                                212/71
CGG TGC CTG GGC ATC CAG CAA AAC GGG GAG CTC AAG AAC GAT TCA TGA ACG AGG GGT CGT
arg cys leu gly ile gln gln asn gly glu leu lys asn asp ser OPA thr arg gly arg
242/81                                272/91
CAC CAA CGT CGA AAC CGA CCG TTG CCA GCC GGC CCA CGA TAT TGC GTG CTC GAG GGT CCG
his gln arg arg asn arg arg leu pro ala gly pro arg tyr cys val leu glu gly pro
302/101                               332/111
CTG TAC OCT CAC CGA ACG TGA GTC CCA CAC CGC GGA GGC GGG CGA CTC TGG CGT CGT TAG
leu tyr pro his arg thr OPA val pro his arg gly gly gly arg leu trp arg arg AMB
362/121                               392/131
CAG CCG AGC TCA AGG TGT CCC GCA CCA CTG TCT CGA ATG CTT TTA ACC GAC CGG ATC CAG
gln pro ser ser arg cys pro ala pro leu ser arg met leu leu thr asp arg ile gln
422/141
AAG GAG AAG ATC
lys glu lys ile

```

SEQ ID N° 6

FIGURE 10B

-2A29.96-Seq.17 [3 to 433] -&gt; 1-phase Translation

DNA sequence 433 b.p. GGATCCTGATGC ... AAGGAGAAGATC linear

```

3/1                                33/11
ATC CTG ATG CAA GTG GTC CCG GAT TTG TCG GCA GCC ACG GCG GTC CCG TCG ACC AAC GTT
ile leu met gln val val arg asp leu ser ala ala thr ala val pro ser thr asn val
63/21                                93/31
GGT GCA TCC GGG CTG CGA GCA TGC ACG CAC CGA CCA GCG CCG CGA GCG CCG CTA GCT GCT
gly ala ser gly leu arg ala cys thr his arg pro ala arg arg ala arg leu ala ala
123/41                                153/51
TGC CCA CTG TTC CTC CCT GCC GGC ACC ATG TGC GAC AAG CTT AAG CCG AGC AGT ACC GGC
cys pro leu phe leu pro ala gly thr met cys asp lys leu lys arg ser ser thr gly
183/61                                213/71
GGT GCC TGG GCA TCC AGC AAA ACG GGG AGC TCA AGA ACG ATT CAT GAA CGA GCG GTC GTC
gly ala trp ala ser ser lys thr gly ser ser arg thr ile his glu arg gly val val
243/81                                273/91
ACC AAC GTC GAA ACC GAC GGT TGC CAG CCG GCC CAC GAT ATT GCG TGC TCG AGG GTC CCG
thr asn val glu thr asp gly cys gln pro ala his asp ile ala cys ser arg val arg
303/101                                333/111
TGT ACC CTC ACC GAA CGT GAG TCC CAC ACC GCG GAG GCG GCG GAC TCT GCG GTC GTT AGC
cys thr leu thr glu arg glu ser his thr ala glu ala gly asp ser gly val val ser
363/121                                393/131
AGC CGA GCT CAA GGT GTC CCG CAC CAC TGT CTC GAA TGC TTT TAA CCG ACC GGA TCC AGA
ser arg ala gln gly val pro his his cys leu glu cys phe och pro thr gly ser arg
423/141
AGG AGA AGA TC
arg arg arg

```

SEQ ID N° 6

FIGURE 10C



## 5B5.95-Seq.24 -&gt; 1-phase Translation

DNA sequence 324 b.p. CCGTCGGCAACT ... TCTCACCAGATC linear

```

1/1                                31/11
CCG TCG GCA ACT TGG CCG CTG AGG TCG GCT TGA TCC CTG GGC CGA GGC GGG TCA GCC AAT
pro ser ala thr trp pro leu arg ser ala OPA ser leu gly arg gly gly ser ala asn
61/21                                91/31
AGC GGC TCC ATC GGC TTT GCT GGT AGC GGT TCG GCG GGA AGC TAG CCG CGA CGT TGT CCG
ser gly ser ile gly phe ala gly ser gly ser ala gly ser AMB arg arg arg cys arg
121/41                                151/51
TGG CCG GTG ATA TAT TGG GTC AGA CCG GTA TGG CCG CCG CTG AGG TGA TCT GCG ACA CGC
trp pro val ile tyr trp val arg arg val trp arg arg leu arg OPA ser ala thr arg
181/61                                211/71
CGC CGC GGT GCT CGA GCC AGG CTT ACG ACC AGG GAA TTT CGA AAA TGT TAT TCA GAA CAT
arg arg gly ala arg ala arg leu thr thr arg glu phe arg lys cys tyr ser glu his
241/81                                271/91
CTT GTA TCT CTT CTC CGT GCC ACC CCC TAG GTG TAG TGT TTT CGA GTA CCG GCA GAT CCC
leu val ser leu leu arg ala thr pro AMB val AMB cys phe arg val pro ala asp pro
301/101
AGG TTC ACC AGG TCT CAC CAG ATC
arg phe thr arg ser his gln ile

```

SEQ ID N° 7

FIGURE 11A

## SB5.95-Seq.24 [2 to 324] -&gt; 1-phase Translation

DNA sequence 324 b.p. CCGTCGGCAACT ... TCTCACCAGATC linear

2/1 32/11  
 CGT CGG CAA CTT GGC CGC TGA GGT CGG CTT GAT CCC TGG GCC GAG GCG GGT CAG CCA ATA  
 arg arg gln leu gly arg OPA gly arg leu asp pro trp ala glu ala gly gln pro ile  
 62/21 92/31  
 GCG GCT CCA TCG GCT TTG CTG GTA GCG GTT CCG CGG GAA GCT AGC GCC GAC GTT GTC GGT  
 ala ala pro ser ala leu leu val ala val arg arg glu ala ser gly asp val val gly  
 122/41 152/51  
 GGC CGG TGA TAT ATT GGG TCA GAC GGG TAT GGC GGC GGC TGA GGT GAT CTG CGA CAC GCC  
 gly arg OPA tyr ile gly ser asp gly tyr gly gly gly OPA gly asp leu arg his ala  
 182/61 212/71  
 GCC GCG GTG CTC GAG CCA GGC TTA CGA CCA GCG AAT TTC GAA AAT GTT ATT CAG AAC ATC  
 ala ala val leu glu pro gly leu arg pro gly asn phe glu asn val ile gln asn ile  
 242/81 272/91  
 TTG TAT CTC TTC TCC GTG CCA CCC CCT AGG TGT AGT GTT TTC GAG TAC CCG CAG ATC CCA  
 leu tyr leu phe ser val pro pro pro arg cys ser val phe glu tyr arg gln ile pro  
 302/101  
 GGT TCA CCA GGT CTC ACC AGA TC  
 gly ser pro gly leu thr arg

SEQ ID N° 7

FIGURE 11B

## 5B5.95-Seq.24 [3 to 324] -&gt; 1-phase Translation

DNA sequence 324 b.p. CCGTCGGCAACT ... TCTCACCAGATC linear

```

3/1                                33/11
GTC GGC AAC TTG GCC GCT GAG GTC GGC TTG ATC CCT GGG CCG AGG CCG GTC AGC CAA TAG
val gly asn leu ala ala glu val gly leu ile pro gly pro arg arg val ser gln AMB
63/21                                93/31
CGG CTC CAT CCG CTT TGC TGG TAG CCG TTC GGC GGG AAG CTA GCG GCG ACG TTG TCG GTG
arg leu his arg leu cys trp AMB arg phe gly gly lys leu ala ala thr leu ser val
123/41                                153/51
GCC GGT GAT ATA TTG GGT CAG ACG GGT ATG GCG GCG GCT GAG GTG ATC TGC GAC ACG CCG
ala gly asp ile leu gly gln thr gly met ala ala ala glu val ile cys asp thr pro
183/61                                213/71
CCG CGG TGC TCG AGC CAG GCT TAC GAC CAG GGA ATT TCG AAA ATG TTA TTC AGA ACA TCT
pro arg cys ser ser gln ala tyr asp gln gly ile ser lys met leu phe arg thr ser
243/81                                273/91
TGT ATC TCT TCT CCG TGC CAC CCC CTA GGT GTA GTG TTT TCG AGT ACC GGC AGA TCC CAG
cys ile ser ser pro cys his pro leu gly val val phe ser ser thr gly arg ser gln
303/101
GTT CAC CAG GTC TCA CCA GAT C
val his gln val ser pro asp

```

SEQ ID N° 7

FIGURE 11C

## -1C7.96-Seq.3 -&gt; 1-phase Translation

DNA sequence 432 b.p. CTTTGCCTGATG ... TCGGCGACGATC linear

1/1 31/11  
 CTT TGC GTG ATG TCC AAT GGC GAA AAC GAC GCC TTG TCA TCG CAA TCG TCA GCA CCG GCC  
 leu cys val met ser asn gly glu asn asp ala leu ser ser gln ser ser ala pro ala  
 61/21 91/31  
 TAG TTT TCG CGA TGA CGC TCG TTC TGA CCG GAC TTG TGA ACG GGT TTC GGG TCG AGG CCG  
 AMB phe ser arg OPA arg ser phe OPA pro asp leu OPA thr gly phe gly ser arg pro  
 121/41 151/51  
 AGC GAA CCG TCG ATT CCA TGG GTG TCG ACG CAT TCG TGG TCA AGG CCG GCG CCG CAG GAC  
 ser glu pro ser ile pro trp val ser thr his ser trp ser arg pro ala arg gln asp  
 181/61 211/71  
 CGT TCC TGG GTT CGA CAC CAT TCG CCC AAA TCG ACC TGC CCC AGG TTG CTC GTG CGC CTG  
 arg ser trp val arg his his ser pro lys ser thr cys pro arg leu leu val arg leu  
 241/81 271/91  
 GCG TCT TGG CTG CCG CCG CAC TAG CGA CTG GCG CGT CGA CGA TCC GGC AGG GCA CGT CAG  
 ala ser trp leu pro pro his AMB arg leu arg arg arg ser gly arg ala arg gln  
 301/101 331/111  
 CGC GAA ACG TCA CCG CGT TCG GCG CAC CAG AGC ACG GAC CCG GCA TGC CCG GGG TCT CCG  
 arg glu thr ser pro arg ser gly his gln ser thr asp pro ala cys arg gly ser arg  
 361/121 391/131  
 ACG GTC GGG CGC CAT CGA CCG CCG ACG AGG TCG CCG TGT CGA GCA CCG TCG GCC GAA ACC  
 thr val gly arg his arg arg arg thr arg ser arg cys arg ala arg trp ala glu thr  
 421/141  
 TCG GCG ACG ATC  
 ser ala thr ile

SEQ ID N° 8

FIGURE 12A

23/73

-1C7.96-Seq.3 [2 to 432] -&gt; 1-phase Translation

DNA sequence 432 b.p. CTTTGCGTGATG ... TCGGCGACGATC linear

```

2/1                                32/11
TTT GCG TGA TGT CCA ATG GCG AAA ACG ACG CCT TGT CAT CGC AAT CGT CAG CAC CGG CCT
phe ala OPA cys pro met ala lys thr thr pro cys his arg asn arg gln his arg pro
62/21                                92/31
AGT TTT CGC GAT GAC GCT CGT TCT GAC CGG ACT TGT GAA CGG GTT TCG GGT CGA GGC CGA
ser phe arg asp asp ala arg ser asp arg thr cys glu arg val ser gly arg gly arg
122/41                                152/51
GCG AAC CGT CGA TTC CAT GGG TGT CGA CGC ATT CGT GGT CAA GGC CGG CGC GGC AGC ACC
ala asn arg arg phe his gly cys arg arg ile arg gly gln gly arg arg gly arg thr
182/61                                212/71
GTT CCT GGG TTC GAC ACC ATT CGC CCA AAT CGA CCT GCC CCA GGT TGC TCG TGC GCC TGG
val pro gly phe asp thr ile arg pro asn arg pro ala pro gly cys ser cys ala trp
242/81                                272/91
CGT CTT GGC TGC CGC CCC ACT AGC GAC TGC GCC GTC GAC GAT CCG GCA GGG CAC GTC AGC
arg leu gly cys arg pro thr ser asp cys ala val asp asp pro ala gly his val ser
302/101                               332/111
GCG AAA CGT CAC CGC GTT CGG GGC ACC AGA GCA CGG ACC CGG CAT GCC GCG GGT CTC GGA
ala lys arg his arg val arg gly thr arg ala arg thr arg his ala ala gly leu gly
362/121                               392/131
CGG TCG GGC GCC ATC GAC GCC GGA CGA GGT CGC GGT GTC GAG CAC GCT GGG CCG AAA CCT
arg ser gly ala ile asp ala gly arg gly arg gly val glu his ala gly pro lys pro
422/141
CGG CGA CGA TC
arg arg arg

```

SEQ ID N° 8

FIGURE 12B

24/73

-1C7.96-Seq.3 [3 to 432] -&gt; 1-phase Translation

DNA sequence 432 b.p. CTTTGGCTGATG ... TCGGCGACGATC linear

3/1 33/11  
 TTG CGT GAT GTC CAA TGG CGA AAA CGA CGC CTT GTC ATC GCA ATC GTC AGC ACC GGC CTA  
 leu arg asp val gln trp arg lys arg arg leu val ile ala ile val ser thr gly leu  
 63/21 93/31  
 GTT TTC GCG ATG ACG CTC GTT CTG ACC GGA CTT GTG AAC GGG TTT CCG GTC GAG GGC GAG  
 val phe ala met thr leu val leu thr gly leu val asn gly phe arg val glu ala glu  
 123/41 153/51  
 CGA ACC GTC GAT TCC ATG GGT GTC GAC GCA TTC GTG GTC AAG GGC GGC GCG GCA GGA CCG  
 arg thr val asp ser met gly val asp ala phe val val lys ala gly ala ala gly pro  
 183/61 213/71  
 TTC CTG GGT TCG ACA CCA TTC GCC CAA ATC GAC CTG CCC CAG GTT GCT CGT GCG CCT GGC  
 phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala arg ala pro gly  
 243/81 273/91  
 GTC TTG GCT GCC GCC CCA CTA GCG ACT GCG CCG TCG ACG ATC CCG CAG GGC ACG TCA GCG  
 val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly thr ser ala  
 303/101 333/111  
 CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC GGC ATG CCG CCG GTC TCG GAC  
 arg asn val thr ala phe gly ala pro glu his gly pro gly met pro arg val ser asp  
 363/121 393/131  
 GGT CCG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG AGC ACG CTG GGC CGA AAC CTC  
 gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu gly arg asn leu  
 423/141  
 GGC GAC GAT C  
 gly asp asp

SEQ ID N° 8

FIGURE 12C

25/73

## 2D7.95-Seq.2 -&gt; 1-phase Translation

DNA sequence 320 b.p. CGAGGCCGAGCG ... TCGGCGACGATC linear

```

1/1                                31/11
CGA GGC CGA GCG AAC CGT CGA TTC CAT GGG TGT CGA CGC ATT CGT GGT CAA GGC GCG CGC
arg gly arg ala asn arg arg phe his gly cys arg arg ile arg gly gln gly arg arg
61/21                                91/31
GGC AGG ACC GTT CCT GGG TTC GAC ACC ATT CGC CCA AAT CGA CCT GCC CCA GGT TGC TCG
gly arg thr val pro gly phe asp thr ile arg pro asn arg pro ala pro gly cys ser
121/41                               151/51
TGC GCC TGG CGT CTT GGC TGC CGC CCC ACT AGC GAC TGC GCC GTC GAC GAT CCG GCA GGG
cys ala trp arg leu gly cys arg pro thr ser asp cys ala val asp asp pro ala gly
181/61                               211/71
CAC GTC ACC GCG AAA CGT CAC CGC GTT CGG GGC ACC AGA GCA CCG ACC CCG CAT GCC GCG
his val ser ala lys arg his arg val arg gly thr arg ala arg thr arg his ala ala
241/81                               271/91
GGT CTC GGA CCG TCG GGC GCC ATC GAC GCC GGA CGA GGT CGC GGT GTC GAG CAC GCT GGG
gly leu gly arg ser gly ala ile asp ala gly arg gly arg gly val glu his ala gly
301/101
CCG AAA CCT CCG CGA CGA TC
pro lys pro arg arg arg

```

SEQ ID N° 9

FIGURE 13A

26/73

2D7.95-Seq.2 [2 to 320] -&gt; 1-phase Translation

DNA sequence 320 b.p. CGAGGCCGAGCG ... TCGGCGACGATC linear

2/1 32/11  
 GAG GCC GAG CGA ACC GTC GAT TCC ATG GGT GTC GAC GCA TTC GTG GTC AAG GCC GGC GCG  
 glu ala glu arg thr val asp ser met gly val asp ala phe val val lys ala gly ala  
 62/21 92/31  
 GCA GGA CCG TTC CTG GGT TCG ACA CCA TTC GCC CAA ATC GAC CTG CCC CAG GTT GCT CGT  
 ala gly pro phe leu gly ser thr pro phe ala gln ile asp leu pro gln val ala arg  
 122/41 152/51  
 GCG CCT GGC GTC TTG GCT GCC GGC CCA CTA GCG ACT GCG CCG TCG ACG ATC CCG CAG GGC  
 ala pro gly val leu ala ala ala pro leu ala thr ala pro ser thr ile arg gln gly  
 182/61 212/71  
 ACG TCA CCG CGA AAC GTC ACC GCG TTC GGG GCA CCA GAG CAC GGA CCC GGC ATG CCG CCG  
 thr ser ala arg asn val thr ala phe gly ala pro glu his gly pro gly met pro arg  
 242/81 272/91  
 GTC TCG GAC GGT CCG GCG CCA TCG ACG CCG GAC GAG GTC GCG GTG TCG ACG ACG CTG GGC  
 val ser asp gly arg ala pro ser thr pro asp glu val ala val ser ser thr leu gly  
 302/101  
 CGA AAC CTC GGC GAC GAT C  
 arg asn leu gly asp asp

SEQ ID N° 9

FIGURE 13B



## 2D7.95-Seq.2 [3 to 320] -&gt; 1-phase Translation

DNA sequence 320 b.p. CGAGGCCGAGCG ... TCGGCGACGATC linear

3/1 33/11  
 AGG CCG AGC GAA CCG TCG ATT CCA TGG GTG TCG ACG CAT TCG TGG TCA AGG CCG GCG CCG  
 arg pro ser glu pro ser ile pro trp val ser thr his ser trp ser arg pro ala arg  
 63/21 93/31  
 CAG GAC CGT TCC TGG GTT CGA CAC CAT TCG CCC AAA TCG ACC TGC CCC AGG TTG CTC GTG  
 gln asp arg ser trp val arg his his ser pro lys ser thr cys pro arg leu leu val  
 123/41 153/51  
 CGC CTG CCG TCT TGG CTG CCG CCC CAC TAG CGA CTG CCG CGT CGA CGA TCC GGC AGG GCA  
 arg leu ala ser trp leu pro pro his AMB arg leu arg arg arg arg ser gly arg ala  
 183/61 213/71  
 CGT CAG CCG GAA ACG TCA CCG CGT TCG GGG CAC CAG ACC ACG GAC CCG GCA TGC CCC GGG  
 arg gln arg glu thr ser pro arg ser gly his gln ser thr asp pro ala cys arg gly  
 243/81 273/91  
 TCT CGG ACG GTC GGG CGC CAT CGA CGC CGG ACG AGG TCG CCG TGT CGA GCA CGC TGG GCC  
 ser arg thr val gly arg his arg arg arg thr arg ser arg cys arg ala arg trp ala  
 303/101  
 GAA ACC TCG GCG ACG ATC  
 glu thr ser ala thr ile

SEQ ID N° 9

FIGURE 13C

## -IB7-Seq.21 -&gt; 1-phase Translation

DNA sequence 445 b.p. TTAACGACTCAG ... CGCATGCGGATC linear

```

1/1                               31/11
TTA ACG ACT CAG ACG GAA ACG CTT GAA CCG CGA GGT CGC TCC GGA CAC CAA TTT GAC TCG
leu thr thr gln thr glu thr leu glu pro arg gly arg ser gly his gln phe asp ser
61/21                               91/31
GCT CTT TGG CAA TTG AAG GTG AGC TGC GAG CAG CCG GGT GAC CGC ATC GTT GGC CTT GCC
ala leu trp gln leu lys val ser cys glu gln pro gly asp arg ile val gly leu ala
121/41                               151/51
ATC AAT CGC CGG CTC GCG GAC GTA GAT AAT CAG CTC ACC GTT GGG ACC GAC CTC GAC CAG
ile asn arg arg leu ala asp val asp asn gln leu thr val gly thr asp leu asp gln
181/61                               211/71
GGG TCC TTT GTG ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA GTC GGT CAT CGC CTA AGG
gly ser phe val thr ala gly leu asp ala asp asp his arg val gly his arg leu arg
241/81                               271/91
CTA CCG TTC TGA CCT GGG GCT GCG TGG GCG CCG ACG ACG TGA GGC ACG TCA TGT CTC AGC
leu pro phe OPA pro gly ala ala trp ala pro thr thr OPA gly thr ser cys leu ser
301/101                               331/111
GGC CCA CCG CCA CCT CGG TCG CCG GCA GTA TGT CAG CAT GTG CAG ATG ACT CCA CGC AGC
gly pro pro pro pro arg ser pro ala val cys gln his val gln met thr pro arg ser
361/121                               391/131
CTT GTT CGC ATC GTT GGT GTC GTG GTT GCG ACG ACC TTG GCG CTG GTG AGC GCA CCC GGC
leu val arg ile val gly val val val ala thr thr leu ala leu val ser ala pro ala
421/141
GGC GGT CGT GCC GCG CAT GCG GAT C
gly gly arg ala ala his ala asp

```

SEQ ID N° 10

FIGURE 14A

## -1B7-Seq.21 [2 to 445] -&gt; 1-phase Translation

DNA sequence 445 b.p. TTAACGACTCAG ... CGCATGCCGATC linear

```

2/1                               32/11
TAA CGA CTC AGA CGG AAA CGC TTG AAC CGC GAG GTC GCT CCG GAC ACC AAT TTG ACT CGG
OCH arg leu arg arg lys arg leu asn arg glu val ala pro asp thr asn leu thr arg
62/21                               92/31
CTC TTT GGC AAT TGA AGG TGA GCT GCG AGC AGC CGG GTG ACC GCA TCG TTG GCC TTG CCA
leu phe gly asn OPA arg OPA ala ala ser ser arg val thr ala ser leu ala leu pro
122/41                               152/51
TCA ATC GCC GGC TCG CGG ACG TAG ATA ATC AGC TCA CCG TTG GGA CCG ACC TCG ACC AGG
ser ile ala gly ser arg thr AMB ile ile ser ser pro leu gly pro thr ser thr arg
182/61                               212/71
GGT CCT TTG TGA CTG CCG GGC TTG ACG CGG ACG ACC ACA GAG TCG GTC ATC GCC TAA GGC
gly pro leu OPA leu pro gly leu thr arg thr thr thr glu ser val ile ala OCH gly
242/81                               272/91
TAC CGT TCT GAC CTG GGG CTG CGT GGG CGC CGA CGA CGT GAG GCA CGT CAT GTC TCA CGG
tyr arg ser asp leu gly leu arg gly arg arg arg arg glu ala arg his val ser ala
302/101                               332/111
GCC CAC CGC CAC CTC GGT CGC CGG CAG TAT GTC AGC ATG TGC AGA TGA CTC CAC GCA GCC
ala his arg his leu gly arg arg gln tyr val ser met cys arg OPA leu his ala ala
362/121                               392/131
TTG TTC GCA TCG TTG GTG TCG TGG TTG CGA CGA CCT TGG CGC TGG TGA GCG CAC CCG CCG
leu phe ala ser leu val ser trp leu arg arg pro trp arg trp OPA ala his pro pro
422/141
GCG GTC GTG CCG CGC ATG CGG ATC
ala val val pro arg met arg ile

```

SEQ ID N° 10

FIGURE 14B

## -1B7-Seq.21 [3 to 445] -&gt; 1-phase Translation

DNA sequence 445 b.p. TTAACGACTCAG ... CGCATGCGGATC linear

3/1 33/11  
 AAC GAC TCA GAC GGA AAC GCT TGA ACC GCG AGG TCG CTC CGG ACA CCA ATT TGA CTC GGC  
 asn asp ser asp gly asn ala OPA thr ala arg ser leu arg thr pro ile OPA leu gly  
 63/21 93/31  
 TCT TTG GCA ATT GAA GGT GAG CTG CGA GCA GCC GGG TGA CCG CAT CGT TGG OCT TGC CAT  
 ser leu ala ile glu gly glu leu arg ala ala gly OPA pro his arg trp pro cys his  
 123/41 153/51  
 CAA TCG CCG GCT CGC GGA CGT AGA TAA TCA GCT CAC CGT TGG GAC CGA CCT CGA CCA GGG  
 gln ser pro ala arg gly arg arg OCH ser ala his arg trp asp arg pro arg pro gly  
 183/61 213/71  
 GTC CTT TGT GAC TGC CGG GCT TGA CGC GGA CGA CCA CAG AGT CCG TCA TCG OCT AAG GCT  
 val leu cys asp cys arg ala OPA arg gly arg pro gln ser arg ser ser pro lys ala  
 243/81 273/91  
 ACC GTT CTG ACC TGG GGC TGC GTG GGC GCC GAC GAC GTG AGG CAC GTC ATG TCT CAG CCG  
 thr val leu thr trp gly cys val gly ala asp asp val arg his val met ser gln arg  
 303/101 333/111  
 CCC ACC GCC ACC TCG GTC GCC GGC AGT ATG TCA GCA TGT GCA GAT GAC TCC ACG CAG CCT  
 pro thr ala thr ser val ala gly ser met ser ala cys ala asp asp ser thr gln pro  
 363/121 393/131  
 TGT TCG CAT CGT TGG TGT CGT GGT TGC GAC GAC CTT GGC GCT GGT GAG CGC ACC CGC CGG  
 cys ser his arg trp cys arg gly cys asp asp leu gly ala gly glu arg thr arg arg  
 423/141  
 CGG TCG TGC CGC GCA TGC GGA TC  
 arg ser cys arg ala cys gly

SEQ ID N° 10

FIGURE 14C

31/73

## DP1034-Seq. 18 -&gt; 1-phase Translation

DNA sequence 510 b.p. CCCGAAGAGGTC ... CGTGACCAGATC linear

1/1 31/11  
 CCC GAA GAG GTC CCC CGT TTT GTT AAT TTT TAA AAA ATT TGT GTC ACA AAG CGG GGT ACC  
 pro glu glu val pro arg phe val asn phe OCH lys ile cys val thr lys arg gly thr  
 61/21 91/31  
 AAG GCA TAA AAC CTA GTA CCT GGG GCG GCG GAT TCA ACG AAA ACC GAG TGG GGG TAG TCA  
 lys ala OCH asn leu val pro gly ala ala asp ser thr lys thr glu trp gly AMB ser  
 121/41 151/51  
 GGG GCG TGC ATT CCG ACG ACC CTG TAC GAC CCG CTG GTG GCA ACG CCG ATG AGT GCG CCG  
 gly ala cys ile pro thr thr leu tyr asp pro leu val ala thr pro met ser ala pro  
 181/61 211/71  
 ACG AAG GCC GAG CGA CGG GCT GCC GGC GCT GAC CGC CGC GGA AGC CGC CGA GTG GAT GGT  
 thr lys ala glu arg arg ala ala gly ala asp arg arg gly ser arg arg val asp gly  
 241/81 271/91  
 CAC CAC CGC CCG CAC CCG ACC GGT ACG GAT CGC GCC TCG GGT TAC CGT CGC CGT CAA CGC  
 his his arg pro his pro thr gly thr asp arg ala ser gly tyr arg arg arg gln arg  
 301/101 331/111  
 GCT GGA CAG CAT CCG TCC CCG CTG GGT CAA TGC ACT CAT GCA GCG CCG CAA CGA ACA GCT  
 ala gly gln his arg ser pro leu gly gln cys thr his ala ala pro gln arg thr ala  
 361/121 391/131  
 CAA CCC TTG AAC CGG GTC CCG GCC TGC CGA CCC TCG GCC GCC GGC CTG CCG CTA CGT GAT  
 gln pro leu asn arg val pro ala cys arg pro ser ala ala gly val pro leu arg asp  
 421/141 451/151  
 AGA CAC AGG GCC ATG GAA ATC CTG GCC AGC CCG ATG CTA CTT CCG CCG GCG GAC TAT CAG  
 arg his arg ala met glu ile leu ala ser arg met leu leu arg pro ala asp tyr gln  
 481/161  
 CGG TCG CTG AGC TTC TAC CGT GAC CAG ATC  
 arg ser leu ser phe tyr arg asp gln ile

SEQ ID N° 11

FIGURE 15A

DP1034-Seq. 18 [2 to 510] -&gt; 1-phase Translation

DNA sequence 510 b.p. CCCGAAGAGGTC ... CGTGACCAGATC linear

```

2/1                               32/11
COG AAG AGG TCC CCC GTT TTG TTA ATT TTT AAA AAA TTT GTG TCA CAA AGC GGG GTA OCA
pro lys arg ser pro val leu leu ile phe lys lys phe val ser gln ser gly val pro
62/21                               92/31
AGG CAT AAA ACC TAG TAC CTG GGG CGG CGG ATT CAA CGA AAA CCG AGT GGG GGT AGT CAG
arg his lys thr AMB tyr leu gly arg arg ile gln arg lys pro ser gly gly ser gln
122/41                               152/51
GGG CGT GCA TTC CGA CGA CCC TGT ACG ACC CGC TGG TGG CAA CGC CGA TGA GTG CGC CGA
gly arg ala phe arg arg pro cys thr thr arg trp trp gln arg arg OPA val arg arg
182/61                               212/71
CGA AGG CCG AGC GAC GGG CTG CCG GCG CTG ACC GCC GCG GAA GCC GCC GAG TGG ATG GTC
arg arg pro ser asp gly leu pro ala leu thr ala ala glu ala ala glu trp met val
242/81                               272/91
ACC ACC GCC CGC ACC CGA CCG GTA CCG ATC CGC OCT CGG GTT ACC GTC GCC GTC AAC GCC
thr thr ala arg thr arg pro val arg ile ala pro arg val thr val ala val asn ala
302/101                               332/111
CTG GAC AGC ATC GGT CCC CGC TGG GTC AAT GCA CTC ATG CAG CGC CGC AAC GAA CAG CTC
leu asp ser ile gly pro arg trp val asn ala leu met gln arg arg asn glu gln leu
362/121                               392/131
AAC CCT TGA ACC GGG TCC CGG CCT GCC GAC OCT CGG CCG CCG GCG TGC CGC TAC GTG ATA
asn pro OPA thr gly ser arg pro ala asp pro arg pro pro ala cys arg tyr val ile
422/141                               452/151
GAC ACA GGG CCA TGG AAA TCC TGG CCA GCC GGA TGC TAC TTC GCC CCG CGG ACT ATC AGC
asp thr gly pro trp lys ser trp pro ala gly cys tyr phe gly arg arg thr ile ser
482/161
GGT CGC TGA GCT TCT ACC GTG ACC AGA TC
gly arg OPA ala ser thr val thr arg

```

SEQ ID N° 11

FIGURE 15B

33/73

DP1034-Seq. 18 [3 to .510] -&gt; 1-phase Translation

DNA sequence 510 b.p. CCGAAGAGGTC ... CGTGACCAGATC linear

3/1 33/11  
 CGA AGA GGT CCC CCG TTT TGT TAA TTT TTA AAA AAT TTG TGT CAC AAA GCG GGG TAC CAA  
 arg arg gly pro pro phe cys OCH phe leu lys asn leu cys his lys ala gly tyr gln  
 63/21 93/31  
 GGC ATA AAA CCT AGT ACC TGG GGC GGC GGA TTC AAC GAA AAC CGA GTG GGG GTA GTC AGG  
 gly ile lys pro ser thr trp gly gly gly phe asn glu asn arg val gly val val arg  
 123/41 153/51  
 GGC GTG CAT TCC GAC GAC CCT GTA CGA CCC GCT GGT GGC AAC GGC GAT GAG TGC GCC GAC  
 gly val his ser asp asp pro val arg pro ala gly gly asn ala asp glu cys ala asp  
 183/61 213/71  
 GAA GGC CGA GCG ACG GGC TGC CGG CGC TGA CCG CCG CCG AAG CCG CCG AGT GGA TGG TCA  
 glu gly arg ala thr gly cys arg arg OPA pro pro arg lys pro pro ser gly trp ser  
 243/81 273/91  
 CCA CCG CCC GCA CCC GAC CGG TAC GGA TCG CGC CTC GGG TTA CCG TCG CCG TCA ACG CGC  
 pro pro pro ala pro asp arg tyr gly ser arg leu gly leu pro ser pro ser thr arg  
 303/101 333/111  
 TGG ACA GCA TCG GTC CCC GCT GGG TCA ATG CAC TCA TGC AGC GCC GCA ACG AAC AGC TCA  
 trp thr ala ser val pro ala gly ser met his ser cys ser ala ala thr asn ser ser  
 363/121 393/131  
 ACC CTT GAA CCG GGT CCC GGC CTG CCG ACC CTC GGC CGC CGG CGT GCC GCT ACG TGA TAG  
 thr leu glu pro gly pro gly leu pro thr leu gly arg arg arg ala ala thr OPA AMB  
 423/141 453/151  
 ACA CAG GGC CAT GGA AAT CCT GGC CAG CCG GAT GCT ACT TCG GCC GGC GGA CTA TCA GCG  
 thr gln gly his gly asn pro gly gln pro asp ala thr ser ala gly gly leu ser ala  
 483/161  
 GTC GCT GAG CTT CTA CCG TGA CCA GAT C  
 val ala glu leu leu pro OPA pro asp

SEQ ID N° 11

FIGURE 15C

## -2A23-Seq. 15 -&gt; 1-phase Translation

DNA sequence 316 b.p. GACCGAAGGGAT ... TCCGGAGTGATC linear

```

1/1                               31/11
GAC CGA AGG GAT TTC GCG ACT AAC TCG GCC TGT AAG GCA ACG CGA GGT CTT CAT GCC GAG
asp arg arg asp phe ala thr asn ser ala cys lys ala thr arg gly leu his ala glu
61/21                               91/31
GAC GTA GAC AGG AAG AGA CAG GGA AGC TGA TGA CGT CGC GTA CCG GAC CGC CAT TCT GTC
asp val asp arg lys arg gln gly ser OPA OPA arg arg val pro asp arg his ser val
121/41                               151/51
GAG TCT TTC CGA GTT CAG CAA CAA TCG ACA CAG AAG CGG GGA CCA GAC CGG GAG GAC GAC
glu ser phe arg val gln gln gln ser thr gln lys arg gly pro asp arg glu asp asp
181/61                               211/71
GCG GCC CGG GCC GCT TCG GGC CGA GTG TCT GAG TAA GAC CAG AGT CAC GGG TCC GTG TGT
ala ala arg ala ala ser gly arg val ser glu OCH asp gln ser his gly ser val cys
241/81                               271/91
GAC AAC CGC GCG GAA TTC AAT CGG ATG GCG GCC GGG ACC GGA TTG CGC CGG TCA CCG AGG
asp asn arg ala glu phe asn arg met ala gly gly thr gly leu arg arg ser pro arg
301/101
AAC CTC CGG AGT GAT C
asn leu arg ser asp

```

SEQ ID N° 12

FIGURE 16A



-2A23-Seq. 15 [2 to 316] -&gt; 1-phase Translation

DNA sequence 316 b.p. GACCGAAGGGAT ... TCCGGAGTGATC linear

2/1	32/11
ACC GAA GGG ATT TCG CGA CTA ACT CGG CCT GTA AGG CAA CGC GAG GTC TTC ATG CCG AGG	
thr glu gly ile ser arg leu thr arg pro val arg gln arg glu val phe met pro arg	
62/21	92/31
ACG TAG ACA GGA AGA GAC AGG GAA GCT GAT GAC GTC GCG TAC CGG ACC GCC ATT CTG TCG	
thr AMB thr gly arg asp arg glu ala asp asp val ala tyr arg thr ala ile leu ser	
122/41	152/51
AGT CTT TCC GAG TTC AGC AAC AAT CGA CAC AGA AGC GGG GAC CAG ACC GGG AGG ACG ACG	
ser leu ser glu phe ser asn asn arg his arg ser gly asp gln thr gly arg thr thr	
182/61	212/71
CGG CCC GGG CCG CTT CGG GCC GAG TGT CTG AGT AAG ACC AGA GTC ACG GGT CCG TGT GTG	
arg pro gly pro leu arg ala glu cys leu ser lys thr arg val thr gly pro cys val	
242/81	272/91
ACA ACC GCG CGG AAT TCA ATC GGA TGG CGG GCG GGA CCG GAT TGC GCC GGT CAC CGA GGA	
thr thr ala arg asn ser ile gly trp arg ala gly pro asp cys ala gly his arg gly	
302/101	
ACC TCC GGA GTG ATC	
thr ser gly val ile	

SEQ ID N° 12

FIGURE 16B

36/73

## -2A23-Seq. 15 [3 to 316] -&gt; 1-phase Translation

DNA sequence 316 b.p. GACCGAAGGGAT ... TCOGGAGTGATC linear

3/1

CCG AAG GGA TTT CCG GAC TAA CTC GGC CTG TAA GGC AAC GCG AGG TCT TCA TGC CGA GGA  
 pro lys gly phe arg asp OCH leu gly leu OCH gly asn ala arg ser ser cys arg gly  
 63/21

33/11

CGT AGA CAG GAA GAG ACA GGG AAG CTG ATG ACG TCG CGT ACC GGA CCG CCA TTC TGT CGA  
 arg arg gln glu glu thr gly lys leu met thr ser arg thr gly pro pro phe cys arg  
 123/41

153/51

GTC TTT CCG AGT TCA GCA ACA ATC GAC ACA GAA GCG GGG ACC AGA CCG GGA GGA CGA CCG  
 val phe pro ser ser ala thr ile asp thr glu ala gly thr arg pro gly gly arg arg  
 183/61

213/71

GGC CCG GGC CCG TTC GGG CCG AGT GTC TGA GTA AGA CCA GAG TCA CCG GTC CGT GTC TGA  
 gly pro gly arg phe gly pro ser val OPA val arg pro glu ser arg val arg val OPA  
 243/81

273/91

CAA CCG CCG GGA ATT CAA TCG GAT GGC GGG CCG GAC CCG ATT GCG CCG GTC ACC GAG GAA  
 gln pro arg gly ile gln ser asp gly gly arg asp arg ile ala pro val thr glu glu  
 303/101

CCT CCG GAG TGA TC  
 pro pro glu OPA

SEQ ID N° 12

FIGURE 16C

## -IC1.96-Seq.20 -&gt; 1-phase Translation

DNA sequence 703 b.p. GGGATTTCGTG ... CCATGCGGATC linear

```

1/1                               31/11
GGG ATT TCG TTG CCC GAT GGA TTG TTT GTA CCG TTT GGG AAA AAC ACT TGA AGT OCT TTT
gly ile ser leu pro asp gly leu phe val arg phe gly lys asn thr OPA ser pro phe
61/21                               91/31
TAT TGG CAA TGC TGG AAA TGG ACA TTC CAA TAT TGC GCG AAT TAA CCG AAC ACG GTG AGG
tyr trp gln cys trp lys trp thr phe gln tyr cys ala asn OCH pro asn thr val arg
121/41                               151/51
GGG GGG CAA GCG TTT GTA CCG GGG CCA GCA AGC GCC GCG GAC CCG TTG ACC GAA GCC AGC
gly gly gln ala phe val pro gly pro ala ser ala ala asp arg leu thr glu ala ser
181/61                               211/71
ATG TTG TTG TGT CAG CGC GGG CTT GGT CTC GAT GTC CCG GCC TTG GCT GGA CCC GCT TCT
met leu leu cys gln arg gly leu gly leu asp val pro ala leu ala gly pro ala ser
241/81                               271/91
TCA AAA CAG GTT GAA CTT AAC GAC TCA AGA ACG GAA ACG CTT GAA CCG CGA CGT CGC TCC
ser lys gln val glu leu asn asp ser arg thr glu thr leu glu pro arg arg arg ser
301/101                               331/111
GGA CAC CAA TTT GAC TCG GCT CTT TGG CAA TTG AAG GTG AGC TGC GAG CAG CCG GGT GAC
gly his gln phe asp ser ala leu trp gln leu lys val ser cys glu gln pro gly asp
361/121                               391/131
CGC ATC GTT GGC CTT GCC ATC AAT CGC CGG CTC GCG GAC GTA GAT AAT CAG CTC ACC GTT
arg ile val gly leu ala ile asn arg arg leu ala asp val asp asn gln leu thr val
421/141                               451/151
GGG ACC GAC CTC GAC CAG CGG TCC TTT GTG ACT GCC GGG CTT GAC GCG GAC GAC CAC AGA
gly thr asp leu asp gln gly ser phe val thr ala gly leu asp ala asp asp his arg
481/161                               511/171
GTC GGT CAT CGC CTA AGG CTA CCG TTC TGA OCT GGG GCT GCG TGG GCG CCG ACG ACG TGA
val gly his arg leu arg leu pro phe OPA pro gly ala ala trp ala pro thr thr OPA
541/181                               571/191
GGC ACG TCA TGT CTC AGC GGC CCA CCG CCA OCT CCG TCG CCG GCA GTA TGT CAG CAT GTG
gly thr ser cys leu ser gly pro pro pro pro arg ser pro ala val cys gln his val
601/201                               631/211
CAG ATG ACT CCA CGC AGC CTT GTT CGC ATC GTT GGT GTC GTG GTT CCG ACG ACC TTG GCG
gln met thr pro arg ser leu val arg ile val gly val val val ala thr thr leu ala
661/221                               691/231
CTG GTG AGC GCA CCC GCC GGC GGT CGT GCC GCG CAT GCG GAT C
leu val ser ala pro ala gly gly arg ala ala his ala asp

```

SEQ ID N° 13

FIGURE 17A

-IC1.96-Seq.20 [2 to 703] -&gt; 1-phase Translation

DNA sequence 703 b.p. GGGATTTCGTG ... CGCATGCCGATC linear

2/1  
 GGA TTT CGT TGC CCG ATG GAT TGT TTG TAC GGT TTG GGA AAA ACA CTT GAA GTC CTT TTT  
 gly phe arg cys pro met asp cys leu tyr gly leu gly lys thr leu glu val leu phe  
 62/21  
 ATT GGC AAT GCT GGA AAT GGA CAT TCC AAT ATT GCG CGA ATT AAC CGA ACA CCG TGA GGG  
 ile gly asn ala gly asn gly his ser asn ile ala arg ile asn arg thr arg OPA gly  
 122/41  
 GGG GGC AAG CGT TTG TAC CCG GGC CAG CAA GCG CCG CCG ACC GGT TGA CCG AAG CCA GCA  
 gly gly lys arg leu tyr arg gly gln gln ala pro pro thr gly OPA pro lys pro ala  
 182/61  
 TGT TGT TGT GTC ACC GCG GGC TTG GTC TCG ATT TCC CCG CCT TGG CTG GAC CCG CTT CTT  
 cys cys cys val ser ala gly leu val ser met ser arg pro trp leu asp pro leu leu  
 242/81  
 CAA AAC AGG TTG AAC TTA ACG ACT CAA GAA CCG AAA CCG TTG AAC CCG GAC GTC GCT CCG  
 gln asn arg leu asn leu thr thr gln glu arg lys arg leu asn arg asp val ala pro  
 302/101  
 GAC ACC AAT TTG ACT CCG CTC TTT GGC AAT TGA AGG TGA GCT GCG AGC AGC CCG GTG ACC  
 asp thr asn leu thr arg leu phe gly asn OPA arg OPA ala ala ser ser arg val thr  
 362/121  
 GCA TCG TTG GCC TTG CCA TCA ATC GCC GGC TCG CCG ACG TAG ATA ATC AGC TCA CCG TTG  
 ala ser leu ala leu pro ser ile ala gly ser arg thr AMB ile ile ser ser pro leu  
 422/141  
 GGA CCG ACC TCG ACC AGG GGT CCT TTG TGA CTG CCG GGC TTG ACG CCG ACG ACC ACA GAG  
 gly pro thr ser thr arg gly pro leu OPA leu pro gly leu thr arg thr thr thr glu  
 482/161  
 TCG GTC ATC GCC TAA GGC TAC CGT TCT GAC CTG GGG CTG CGT GCG CCG CGA CGA CGT GAG  
 ser val ile ala OCH gly tyr arg ser asp leu gly leu arg gly arg arg arg arg glu  
 542/181  
 GCA CGT CAT GTC TCA GCG GCC CAC CCG CAC CTC GGT CCG CCG CAG TAT GTC AGC ATG TGC  
 ala arg his val ser ala ala his arg his leu gly arg arg gln tyr val ser met cys  
 602/201  
 AGA TGA CTC CAC GCA GCC TTG TTC GCA TCG TTG GTG TCG TGG TTG CGA CGA CCT TGG CCG  
 arg OPA leu his ala ala leu phe ala ser leu val ser trp leu arg arg pro trp arg  
 662/221  
 TGG TGA GCG CAC CCG CCG GCG GTC GTG CCG CGC ATG CCG ATC  
 trp OPA ala his pro pro ala val val pro arg met arg ile

SEQ ID N° 13

FIGURE 17B

-1C1.96-Seq.20 [3 to 703] -&gt; 1-phase Translation

DNA sequence 703 b.p. GGGATTTCGTTG ... CGCATGCGGATC linear

```

3/1                                33/11
GAT TTC GTT GCC CGA TGG ATT GTT TGT ACG GTT TGG GAA AAA CAC TTG AAG TCC TTT TTA
asp phe val ala arg trp ile val cys thr val trp glu lys his leu lys ser phe leu
63/21                                93/31
TTG GCA ATG CTG GAA ATG GAC ATT CCA ATA TTG CGC GAA TTA ACC GAA CAC CGT GAG GGG
leu ala met leu glu met asp ile pro ile leu arg glu leu thr glu his gly glu gly
123/41                                153/51
GGG GCA AGC GTT TGT ACC GGG GCC AGC AAG CGC CGC CGA CCG GTT GAC CGA AGC CAG CAT
gly ala ser val cys thr gly ala ser lys arg arg arg pro val asp arg ser gln his
183/61                                213/71
GTT GTT GTG TCA GCG CGG GCT TGG TCT CGA TGT CCC GGC CTT GCC TCG ACC CGC TTC TTC
val val val ser ala arg ala trp ser arg cys pro gly leu gly trp thr arg phe phe
243/81                                273/91
AAA ACA GGT TGA ACT TAA CGA CTC AAG AAC GGA AAC GCT TGA ACC GCG ACG TCG CTC CGG
lys thr gly OPA thr OCH arg leu lys asn gly asn ala OPA thr ala thr ser leu arg
303/101                                333/111
ACA CCA ATT TGA CTC GGC TCT TTG GCA ATT GAA GGT GAG CTG CGA GCA GCC GGG TGA CCG
thr pro ile OPA leu gly ser leu ala ile glu gly glu leu arg ala ala gly OPA pro
363/121                                393/131
CAT CGT TGG CCT TGC CAT CAA TCG CCG GCT CGC GGA CGT AGA TAA TCA GCT CAC CGT TGG
his arg trp pro cys his gln ser pro ala arg gly arg arg OCH ser ala his arg trp
423/141                                453/151
GAC CGA CCT CGA CCA GGG GTC CTT TGT GAC TGC CGG GCT TGA CGC GCA CGA CCA CAG AGT
asp arg pro arg pro gly val leu cys asp cys arg ala OPA arg gly arg pro gln ser
483/161                                513/171
CGG TCA TCG CCT AAG GCT ACC GTT CTG ACC TGG GGC TGC CTG GCC CCC GAC GAC GTG AGG
arg ser ser pro lys ala thr val leu thr trp gly cys val gly ala asp asp val arg
543/181                                573/191
CAC GTC ATG TCT CAG CGG CCC ACC GCC ACC TCG GTC GCC GGC AGT ATG TCA GCA TGT GCA
his val met ser gln arg pro thr ala thr ser val ala gly ser met ser ala cys ala
603/201                                633/211
GAT GAC TCC ACG CAG CCT TGT TCG CAT CGT TGG TGT CGT GGT TCC GAC GAC CTT GGC GCT
asp asp ser thr gln pro cys ser his arg trp cys arg gly cys asp asp leu gly ala
663/221                                693/231
GGT GAG CGC ACC CGC CGG CGG TCG TGC CGC-GCA TCC GCA TC
gly glu arg thr arg arg arg ser cys arg ala cys gly

```

SEQ ID N° 13

FIGURE 17C

## 5C7.95 -&gt; 1-phase Translation

DNA sequence 361 b.p. CCACCGGGGCTG ... TCTCACCAGATC linear

1/1	31/11
CCA CCG GGG CTG GAG GGG CGA ATG TGC GCC	GAA CGC CGT CGG CCA ACT TGG CCG CTG AGG
pro pro gly leu glu gly arg met cys ala	glu arg arg arg pro thr trp pro leu arg
61/21	91/31
GCG GCT GAT CCC CTG GCC CGA GAC GGG GCA	AGC CAA TAG CGG CTC CAT CGG GCT TTG CTG
ala ala asp pro leu ala arg asp gly ala	ser gln AMB arg leu his arg ala leu leu
121/41	151/51
GTA GCG GTT CCG CGG GAA CCG AGC GCC GAC	GTT GTC GGT GCC CGG TGA TAT ATT GCG TCA
val ala val arg arg glu pro ser ala asp	val val gly ala arg OPA tyr ile gly ser
181/61	211/71
GAC GGG TAT GGC GGC GAC TGA GGT GAT CTG	CGA CAC GCC GCC GCG GTG CTC GAG CCA GGC
asp gly tyr gly gly asp OPA gly asp leu	arg his ala ala ala val leu glu pro gly
241/81	271/91
TTA CGA CCA GGG AAT TTC GAA AAT GTT ATT	CAG AAC ATC TTG TAT CTC TTC CTC CGT GCC
leu arg pro gly asn phe glu asn val ile	gln asn ile leu tyr leu phe leu arg ala
301/101	331/111
ACC CCC TAG GTG TAG TGT TTT CGA GTA CCG	GCA GAT CCC AGT TCA CCA GTC TCA CCA GAT
thr pro AMB val AMB cys phe arg val pro	ala asp pro ser ser pro val ser pro asp

c

SEQ ID N° 14

FIGURE 18A

5C7.95 [2 to 361] -&gt; 1-phase Translation

DNA sequence 361 b.p. CCACCGGGCTG ... TCTCACCAGATC linear

2/1  
 CAC CGG GGC TGG AGG GGC GAA TGT GCG CCG AAC GCC GTC GGC CAA CTT GGC CGC TGA GGG  
 his arg gly trp arg gly glu cys ala pro asn ala val gly gln leu gly arg OPA gly  
 62/21  
 CGG CTG ATC CCC TGG CCC GAG ACG GGG CAA GCC AAT AGC GGC TCC ATC GGG CTT TGC TGG  
 arg leu ile pro trp pro glu thr gly gln ala asn ser gly ser ile gly leu cys trp  
 122/41  
 TAG CGG TTC GGC GGG AAC CGA GCG CCG ACG TTG TCG GTG CCC GGT GAT ATA TTG GGT CAG  
 AMB arg phe gly gly asn arg ala pro thr leu ser val pro gly asp ile leu gly gln  
 182/61  
 ACG GGT ATG GCG GCG ACT GAG GTG ATC TGC GAC ACG CCG CCG CCG TGC TCG AGC CAG GCT  
 thr gly met ala ala thr glu val ile cys asp thr pro pro arg cys ser ser gln ala  
 242/81  
 TAC GAC CAG GGA ATT TCG AAA ATG TTA TTC AGA ACA TCT TGT ATC TCT TCC TCC GTG CCA  
 tyr asp gln gly ile ser lys met leu phe arg thr ser cys ile ser ser ser val pro  
 302/101  
 CCC CCT AGG TGT AGT GTT TTC GAG TAC CGG CAG ATC CCA GTT CAC CAG TCT CAC CAG ATC  
 pro pro arg cys ser val phe glu tyr arg gln ile pro val his gln ser his gln ile  
 332/111

SEQ ID N° 14

FIGURE 18B

5C7.95 [3 to 361] -&gt; 1-phase Translation

DNA sequence 361 b.p. CCACCGGGGCTG ... TCTCACCAGATC linear

3/1	33/11
ACC GGG GCT GGA GGG GCG AAT GTG CGC CGA	ACG CCG TCG GCC AAC TTG GCC CCT GAG GGC
thr gly ala gly gly ala asn val arg arg	thr pro ser ala asn leu ala ala glu gly
63/21	93/31
GGC TGA TCC CCT GGC CCG AGA CGG GGC AAG	CCA ATA GCG GCT CCA TCG GGC TTT GCT GGT
gly OPA ser pro gly pro arg arg gly lys	pro ile ala ala pro ser gly phe ala gly
123/41	153/51
AGC GGT TCG GCG GGA ACC GAG CGC CGA CGT	TGT CCG TGC CCG GTG ATA TAT TGG GTC AGA
ser gly ser ala gly thr glu arg arg arg	cys arg cys pro val ile tyr trp val arg
183/61	213/71
CGG GTA TGG CCG CGA CTG AGG TGA TCT GCG	ACA CGC CGC GGT GCT CGA GCC AGG CTT
arg val trp arg arg leu arg OPA ser ala	thr arg arg arg gly ala arg ala arg leu
243/81	273/91
ACG ACC AGG GAA TTT CGA AAA TGT TAT TCA	GAA CAT CTT GTA TCT CTT CCT CCG TGC CAC
thr thr arg glu phe arg lys cys tyr ser	glu his leu val ser leu pro pro cys his
303/101	333/111
CCC CTA GGT GTA GTG TTT TCG AGT ACC GGC	AGA TCC CAG TTC ACC AGT CTC ACC AGA TC
pro leu gly val val phe ser ser thr gly	arg ser gln phe thr ser leu thr arg

SEQ ID N° 14

FIGURE 18C



## 6C8.95-Seq.19 -&gt; 1-phase Translation

DNA sequence 300 b.p. CAAGCCCGGCCG ... GATACGCGGTAC linear

1/1	31/11
CAA GCC CCG CCG CGA CTG TTT GCC GTT TTG GGG CTC CTA CCA GAA CAC CAC CTG GCG GCC	
gln ala arg pro arg leu phe ala val leu gly leu leu pro glu his his leu ala ala	
61/21	91/31
GCG CAC CAT GGT GTG CAC CAG TTG CGA TCG GTT OCT CCC GCG CCG GGG CCG CGA CGA CGT	
ala his his gly val his gln leu arg ser val pro pro ala arg gly arg arg arg arg	
121/41	151/51
CGA TGC CCG CCG CCC GGC GGC GCA GCT GCG TAG CTC GAC CCG GTC GAC GAC GAC GGG GTC	
arg cys pro arg pro gly gly ala ala ala AMB leu asp pro val asp asp asp gly val	
181/61	211/71
GGC GGA CCA GTC GGC GAT GTC GAG GCG ATG GCA ATA CAG CCG CTT GGT CCG CCG CCA CAC	
gly gly pro val gly asp val glu ala met ala ile gln arg leu gly ala arg pro his	
241/81	271/91
GTC TGA GGT GGC GAA GAC CAG TCC CCG GCC CAC CCG CAG CCG GAT CCG GAT ACC CCG TAC	
val OPA gly gly glu asp gln ser arg ala his arg gln pro asp pro asp thr arg tyr	

SEQ ID N° 15

FIGURE 19A

## 6C8.95-Seq.19 [2 to 300] -&gt; 1-phase Translation

DNA sequence 300 b.p. CAAGCCCGGCCG ... GATACGCGGTAC linear

```

2/1                                32/11
AAG CCC GGC CGC GAC TGT TTG CCG TTT TGG GGC TCC TAC CAG AAC ACC ACC TGG CCG CCG
lys pro gly arg asp cys leu pro phe trp gly ser tyr gln asn thr thr trp arg pro
62/21                                92/31
CGC ACC ATG GTG TGC ACC AGT TGC GAT CCG TTC CTC CCG CGC GCG GGC GGC GAC GAC GTC
arg thr met val cys thr ser cys asp arg phe leu pro arg ala gly gly asp asp val
122/41                                152/51
GAT GCC CGC GGC CCG GCG GCG CAG CTG CGT AGC TCG ACC CGG TCG ACG ACG ACG GCG TCG
asp ala arg ala pro ala ala gln leu arg ser ser thr arg ser thr thr thr gly ser
182/61                                212/71
GCG GAC CAG TCG GCG ATG TCG AGG CGA TGG CAA TAC AGC GCC TTG GTG CCG GGC CAC ACG
ala asp gln ser ala met ser arg arg trp gln tyr ser ala leu val arg gly his thr
242/81                                272/91
TCT GAG GTG GCG AAG ACC AGT CCC GCG CCC ACC GGC AGC CGG ATC CGG ATA CGC GGT AC
ser glu val ala lys thr ser pro ala pro thr gly ser arg ile arg ile arg gly

```

SEQ ID N° 15

FIGURE 19B

## 6C8.95-Seq.19 [3 to 300] -&gt; 1-phase Translation

DNA sequence 300 b.p. CAAGCCCGGCCG ... GATACGCGGTAC linear

3/1	33/11
AGC CCG GCC GCG ACT GTT TGC CGT TTT GGG GCT CCT ACC AGA ACA CCA CCT GGC GGC CGC	
ser pro ala ala thr val cys arg phe gly ala pro thr arg thr pro pro gly gly arg	
63/21	93/31
GCA CCA TGG TGT GCA CCA GTT GCG ATC GGT TCC TCC CGC GCG CGG GCG GCG ACG ACG TCG	
ala pro trp cys ala pro val ala ile gly ser ser arg ala arg ala ala thr thr ser	
123/41	153/51
ATG CCC GCG CCC CGG CGG CGC AGC TGC GTA GCT CGA CCC GGT CGA CGA CGA CGG GGT CGG	
met pro ala pro arg arg arg ser cys val ala arg pro gly arg arg arg arg gly arg	
183/61	213/71
CGG ACC AGT CGG CGA TGT CGA GGC GAT GGC AAT ACA GCG CCT TGG TGC GCG GCC ACA CGT	
arg thr ser arg arg cys arg gly asp gly asn thr ala pro trp cys ala ala thr arg	
243/81	273/91
CTG AGG TGG CGA AGA CCA GTC CCG CGC CCA CCG GCA GCC GGA TCC GGA TAC CCG GTA C	
leu arg trp arg arg pro val pro arg pro pro ala ala gly ser gly tyr ala val	

SEQ ID N° 15

FIGURE 19C

## DE5.strider -&gt; 1-phase Translation

DNA sequence 400 b.p. TGCGCATGCCGA ... CAGTTCGGGATC linear

```

1/1                               31/11
TGC GCA TGC CGA CCA GTG TGG TTG GGC GGA GTT CGT TTG TTC GCG ATT GCC TCA ACG ATT
cys ala cys arg pro val trp leu ala gly val arg leu phe ala ile ala ser thr ile
61/21                               91/31
CGA TAT AAC CAC TCT AGT CAC ATC AAC CAC ACT CGT ACC ATC GAG CGT GTG GGT TCA TGC
arg tyr asn his ser ser his ile asn his thr arg thr ile glu arg val gly ser cys
121/41                               151/51
CAT GCA TTC GCG ACC GCG GGA GGC GGC GAA CCC GGC GGC ACA CAT AAT CCA GAT TGA GGA
his ala phe ala thr ala gly ala gly glu pro gly ala thr his asn pro asp OPA gly
181/61                               211/71
GAC TTC CGT GCG GAA CCG ACG CCG ACG CAA GCT TTC GAC AGC CAT GAG CGC GGT CGC CGC
asp phe arg ala glu pro thr pro thr gln ala phe asp ser his glu arg gly arg arg
241/81                               271/91
OCT GGC AGT TGC AAG TCC TTG TGC ATA TTT TCT TGT CTA CGA ATC AAC CGA AAC GAC CGA
pro gly ser cys lys ser leu cys ile phe ser cys leu arg ile asn arg asn asp arg
301/101                               331/111
CGC GCC CGA GCA CCA TGA ATT CAA GCA GGC GGC GGT GTT GAC CGA CCT GCC CCG CGA GCT
ala ala arg ala pro OPA ile gln ala gly gly gly val asp arg pro ala arg arg ala
361/121                               391/131
GAT GTC CGC GCT ATC CCA GGG GTT GTC CCA GTT CCG GAT C
asp val arg ala ile ala gly val val pro val arg asp

```

SEQ ID N° 16

FIGURE 20A

DE5.strider [2 to 400] -> 1-phase Translation

DNA sequence 400 b.p. TCGCATGCCGA ... CAGTTCGGGATC linear

```

2/1                               32/11
GCG CAT GCC GAC CAG TGT GGT TGG CCG GAG TTC GTT TGT TCG CGA TTG CCT CAA CGA TTC
ala his ala asp gln cys gly trp pro glu phe val cys ser arg leu pro gln arg phe
62/21                               92/31
GAT ATA ACC ACT CTA GTC ACA TCA ACC ACA CTC GTA CCA TCG AGC GTG TGG GTT CAT GCC
asp ile thr thr leu val thr ser thr thr leu val pro ser ser val trp val his ala
122/41                               152/51
ATG CAT TCG CGA CCG CCG GAG CCG GCG AAC CCG GCG CCA CAC ATA ATC CAG ATT GAG GAG
met his ser arg pro arg glu pro ala asn pro ala pro his ile ile gln ile glu glu
182/61                               212/71
ACT TCC GTG CCG AAC CGA CCG CGA CCG AAG CTT TCG ACA GCC ATG AGC GCG GTC GCC GCC
thr ser val pro asn arg arg arg arg lys leu ser thr ala met ser ala val ala ala
242/81                               272/91
CTG GCA GTT GCA AGT CCT TGT GCA TAT TTT CTT GTC TAC GAA TCA ACC GAA ACC ACC GAG
leu ala val ala ser pro cys ala tyr phe leu val tyr glu ser thr glu thr thr glu
302/101                               332/111
CGG CCC GAG CAC CAT GAA TTC AAG CAG GCG GCG GTG TTG ACC GAC CTG CCC GGC GAG CTG
arg pro glu his his glu phe lys gln ala ala val leu thr asp leu pro gly glu leu
362/121                               392/131
ATG TCC GCG CTA TCG CAG GGG TTG TCC CAG TTC GGG ATC
met ser ala leu ser gln gly leu ser gln phe gly ile

```

SEQ ID N° 16

FIGURE 20B

## DE5.strider [3 to 400] -&gt; 1-phase Translation

DNA sequence 400 b.p. TGCGCATGCCGA ... CAGTTCGGGATC linear

3/1 33/11  
 CGC ATG CCG ACC AGT GTG GTT GGC CGG AGT TCG TTT GTT CGC GAT TGC CTC AAC GAT TCG  
 arg met pro thr ser val val gly arg ser ser phe val arg asp cys leu asn asp ser  
 63/21 93/31  
 ATA TAA CCA CTC TAG TCA CAT CAA CCA CAC TCG TAC CAT CGA GCG TGT GGG TTC ATG CCA  
 ile OCH pro leu AMB ser his gln pro his ser tyr his arg ala cys gly phe met pro  
 123/41 153/51  
 TGC ATT CGC GAC CGC GGG AGC CGG CGA ACC CGG CGC CAC ACA TAA TCC AGA TTG AGG AGA  
 cys ile arg asp arg gly ser arg arg thr arg arg his thr OCH ser arg leu arg arg  
 183/61 213/71  
 CTT CCG TGC CGA ACC GAC GCC GAC GCA AGC TTT CGA CAG CCA TGA GCG CGG TCG CCG CCC  
 leu pro cys arg thr asp ala asp ala ser phe arg gln pro OPA ala arg ser pro pro  
 243/81 273/91  
 TGG CAG TTG CAA GTC CTT GTG CAT ATT TTC TTG TCT ACG AAT CAA CCG AAA CGA CCG AGC  
 trp gln leu gln val leu val his ile phe leu ser thr asn gln pro lys arg pro ser  
 303/101 333/111  
 GGC CCG AGC ACC ATG AAT TCA AGC AGG CGG CGG TGT TGA CCG ACC TGC CCG GCG ACC TGA  
 gly pro ser thr met asn ser ser arg arg arg cys OPA pro thr cys pro ala ser OPA  
 363/121 393/131  
 TGT CCG CGC TAT CGC AGG GGT TGT CCC AGT TCG GGA TC  
 cys pro arg tyr arg arg gly cys pro ser ser gly

SEQ ID N° 16

FIGURE 20C

## DE8.strider -&gt; 1-phase Translation

DNA sequence 359 b.p. GCGGGCCACCGA ... CTCGGGCTGATC linear

1/1	31/11
GCG GGC CAC CGA TCA GTC GAT CGG GTG GTT	TCC GCT CCA TCA GCC CGG AAT TGA GGT GCC
ala gly his arg ser val asp arg val val	ser ala pro ser ala arg asn OPA gly ala
61/21	91/31
GCA GTG ACG ACA CCA GCG CAG GAC GCG CCG	TTG GTG TTT CCC TCT GTT GCT TTC CCG TCC
ala val thr thr pro ala gln asp ala pro	leu val phe pro ser val ala phe pro ser
121/41	151/51
GGC TCG OCT TTT TTT CAT CAA CGT TGG ACT	GCC GCA GTG GCG ATG TTG GTC GCC GGC GTG
gly ser pro phe phe his gln arg trp thr	ala ala val ala met leu val ala gly val
181/61	211/71
TTC GGT CAC CTG ACG GTC GGG ATG TTC CTT	GGG TCT CCG GTT GCT GCT GGG TTT GCT CAA
phe gly his leu thr val gly met phe leu	gly ser arg val ala ala gly phe ala gln
241/81	271/91
TGC CCT GCT GGT GCG GCG TTC GGC CGA GTC	GAT CAC CGC CAA AGA GCA CCC GTT AAA ACG
cys pro ala gly ala ala phe gly arg val	asp his arg gln arg ala pro val lys thr
301/101	331/111
GTC GAT GGC OCT CAA CTC GGC ATC GCG ACT	GGC GAT TAT CAC CAT GCC TCG GGC TGA TC
val asp gly pro gln leu gly ile ala thr	gly asp tyr his his ala ser gly OPA

SEQ ID N° 17

FIGURE 21A

## DE8.strider [2 to 359] -&gt; 1-phase Translation

DNA sequence 359 b.p. GCGGGCCACCGA ... CTCGGGCTGATC linear

2/1 32/11  
 CCG GCC ACC GAT CAG TCG ATC GGG TGG TTT CCG CTC CAT CAG CCC GGA ATT GAG GTG CCG  
 arg ala thr asp gln ser ile gly trp phe pro leu his gln pro gly ile glu val pro  
 62/21 92/31  
 CAG TGA CGA CAC CAG CGC AGG ACG CGC CGT TGG TGT TTC OCT CTG TTG CTT TCC CGT CCG  
 gln opa arg his gln arg arg thr arg arg trp cys phe pro leu leu leu ser arg pro  
 122/41 152/51  
 GCT CGC CTT TTT TTC ATC AAC GTT GGA CTG CCG CAG TGG CGA TGT TGG TCG CCG GCG TGT  
 ala arg leu phe phe ile asn val gly leu pro gln trp arg cys trp ser pro ala cys  
 182/61 212/71  
 TCG GTC ACC TGA CGG TCG GGA TGT TCC TTG GGT CTC GGG TTG CTG CTG GGT TTG CTC AAT  
 ser val thr opa arg ser gly cys ser leu gly leu gly leu leu leu gly leu leu asn  
 242/81 272/91  
 GGC CTG CTG GTG CGG CGT TCG GCC GAG TCG ATC ACC GCC AAA GAG CAC CCG TTA AAA CCG  
 ala leu leu val arg arg ser ala glu ser ile thr ala lys glu his pro leu lys arg  
 302/101 332/111  
 TCG ATG GCC CTC AAC TCG GCA TCG CGA CTG GCG ATT ATC ACC ATG CCT CCG GCT GAT C  
 ser met ala leu asn ser ala ser arg leu ala ile ile thr met pro arg ala asp

SEQ ID N° 17

FIGURE 21B



## DE8.strider [3 to 359] -&gt; 1-phase Translation

DNA sequence 359 b.p. GCGGGCCACCGA ... CTGGGCTGATC linear

```
3/1                                33/11
GGG CCA CCG ATC AGT CGA TCG GGT GGT TTC CGC TCC ATC AGC CCG GAA TTG AGG TGC CGC
gly pro pro ile ser arg ser gly gly phe arg ser ile ser pro glu leu arg cys arg
63/21                               93/31
AGT GAC GAC ACC AGC GCA GGA CGC GGC GTT GGT GTT TCC CTC TGT TGC TTT CCC GTC CGG
ser asp asp thr ser ala gly arg ala val gly val ser leu cys cys phe pro val arg
123/41                             153/51
CTC GCC TTT TTT TCA TCA ACG TTG GAC TGC CGC AGT GGC GAT GTT GGT CGC CGG CGT GTT
leu ala phe phe ser ser thr leu asp cys arg ser gly asp val gly arg arg arg val
183/61                             213/71
CGG TCA OCT GAC GGT CGG GAT GTT OCT TGG GTC TCG GGT TGC TCC TCG GTT TGC TCA ATG
arg ser pro asp gly arg asp val pro trp val ser gly cys cys trp val cys ser met
243/81                             273/91
CCC TGC TGG TGC GGC GTT CGG CCG AGT CGA TCA CCG CCA AAG AGC ACC CGT TAA AAC GGT
pro cys trp cys gly val arg pro ser arg ser pro pro lys ser thr arg OCH asn gly
303/101                           333/111
CGA TGG CCC TCA ACT CGG CAT CGC GAC TGG CGA TTA TCA CCA TGC CTC GGG CTG ATC
arg trp pro ser thr arg his arg asp trp arg leu ser pro cys leu gly leu ile
```

SEQ ID N° 17

FIGURE 21C

## DE10.strider -&gt; 1-phase Translation

DNA sequence 390 b.p. GTCGAACAGGTA ... GGTGGGATGATC linear

```

1/1                               31/11
GTC GAA CAG GTA CCG AAG GCG CCG TCG GTC GCT CGG TCC GCT GGT ATC TCG TGT TCA GCC
val glu gln val arg lys ala pro ser val ala arg ser ala gly ile ser cys ser ala
61/21                               91/31
AGC CAG CCG CCG TTA ACG TGG CCG AAC AGG TCG TCT TGG GGT CCG GCA TCA CCG TCG ATG
ser gln arg pro leu thr trp pro asn arg ser ser trp gly arg ala ser ala ser met
121/41                               151/51
TGG CTC AGG TCG ATA CCC GAG GGG ATG GCA AGT GTC ACC CCG CCA TCC TTC CAC CTC TTT
trp leu arg ser ile pro glu gly met ala ser val thr pro pro ser phe his leu phe
181/61                               211/71
TCG GGT GCA ACG ATC GGG CCA TGC CTG ACG GGG AGC AGA GCC AGC CAC CCG CCC AAG AAG
ser gly ala thr ile gly pro cys leu thr gly ser arg ala ser his arg pro lys lys
241/81                               271/91
ATG CCG AAG ACG ACT CCG GGC CCG ACG CCG CCG AGG CCG CCG CCG AAC CCA AAT CAT
met arg lys thr thr arg gly pro thr pro arg arg pro pro arg pro asn pro asn his
301/101                               331/111
CAG CCG GTC CCG ATG TTC TCG ACC TAC GGT ATC GCC TCG ACA CTA CTC GGC GTG CTA TCG
gln pro val pro met phe ser thr tyr gly ile ala ser thr leu leu gly val leu ser
361/121
GTC GCC GCG GTC GTG CTG GGT GCG ATG ATC
val ala ala val val leu gly ala met ile

```

SEQ ID N° 18

FIGURE 22A

## DE10.strider [2 to 390] -&gt; 1-phase Translation

DNA sequence 390 b.p. GTCGAACAGGTA ... GGTGGATGATC linear

2/1 32/11  
 TCG AAC AGG TAC GGA AGG CGC CGT CGG TCG CTC GGT CCG CTG GTA TCT CGT GTT CAG CCA  
 ser asn arg tyr gly arg arg arg arg ser leu gly pro leu val ser arg val gln pro  
 62/21 92/31  
 GCC AGC GGC CGT TAA CGT GGC CGA ACA GGT CGT CTT GGG GTC GGG CAT CAG CGT CGA TGT  
 ala ser gly arg OCH arg gly arg thr gly arg leu gly val gly his gln arg arg cys  
 122/41 152/51  
 GGC TCA GGT CGA TAC CCG AGG GGA TGG CAA GTG TCA CCC CGC CAT CCT TCC ACC TCT TTT  
 gly ser gly arg tyr pro arg gly trp gln val ser pro arg his pro ser thr ser phe  
 182/61 212/71  
 CGG GTG CAA CGA TCG GGC CAT GCC TGA CGG GGA GCA GAG CCA GCC ACC GGC CCA AGA AGA  
 arg val gln arg ser gly his ala OPA arg gly ala glu pro ala thr gly pro arg arg  
 242/81 272/91  
 TGC GGA AGA CGA CTC GCG GCC CGA CGC CGC GGA GGC CGC CGC GGC CGA ACC CAA ATC ATC  
 cys gly arg arg leu ala ala arg arg arg gly gly arg arg gly arg thr gln ile ile  
 302/101 332/111  
 AGC CGG TCC CGA TGT TCT CGA CCT ACG GTA TCG CCT CGA CAC TAC TCG GCG TGC TAT CGG  
 ser arg ser arg cys ser arg pro thr val ser pro arg his tyr ser ala cys tyr arg  
 362/121  
 TCG CCG CGG TCG TGC TGG GTG CGA TGA TC  
 ser pro arg ser cys trp val arg OPA

SEQ ID N° 18

FIGURE 22B

DE10.strider [3 to 390] -&gt; 1-phase Translation

DNA sequence 390 b.p. GTCGAACAGGTA ... GGTGGGATGATC linear

3/1 33/11  
 CGA ACA GGT ACG GAA GGC GGC GTC GGT CGC TCG GTC CGC TGG TAT CTC GTG TTC AGC CAG  
 arg thr gly thr glu gly ala val gly arg ser val arg trp tyr leu val phe ser gln  
 63/21 93/31  
 CCA GCG GGC GTT AAC GTG GCC GAA CAG GTC GTC TTG GGG TCG GGC ATC AGC GTC GAT GTG  
 pro ala ala val asn val ala glu gln val val leu gly ser gly ile ser val asp val  
 123/41 153/51  
 GCT CAG GTC GAT ACC CGA GGG GAT GGC AAG TGT CAC CCC GCC ATC CTT CCA OCT CTT TTC  
 ala gln val asp thr arg gly asp gly lys cys his pro ala ile leu pro pro leu phe  
 183/61 213/71  
 GGG TGC AAC GAT CGG GCC ATG OCT GAC GGG GAG CAG AGC CAG CCA CCG GCC CAA GAA GAT  
 gly cys asn asp arg ala met pro asp gly glu gln ser gln pro pro ala gln glu asp  
 243/81 273/91  
 GCG GAA GAC GAC TCG CCG CCC GAC GCC GCG GAG GCC GCG GCG GOC GAA CCC AAA TCA TCA  
 ala glu asp asp ser arg pro asp ala ala glu ala ala ala glu pro lys ser ser  
 303/101 333/111  
 GCC GGT CCC GAT GTT CTC GAC CTA CCG TAT CGC CTC GAC ACT ACT CCG CGT GCT ATC GGT  
 ala gly pro asp val leu asp leu arg tyr arg leu asp thr thr arg arg ala ile gly  
 363/121  
 CGC CGC GGT CGT GCT GGG TGC GAT GAT C  
 arg arg gly arg ala gly cys asp asp

SEQ ID N° 18

FIGURE 22C

55/73

## DE14.strider -&gt; 1-phase Translation

DNA sequence 380 b.p. GTTCGGCAACGG ... TGCTGGTAGATC linear

```

1/1                               31/11
GTT GCG CAA CGG GGT GAG CAC CGA CGC GAT GAT GGC GCA ACT ATC GAA ACT GCA GGA CAT
val ala gln arg gly glu his arg arg asp asp gly ala thr ile glu thr ala gly his
61/21                               91/31
CGC CAA CGC CAA CGA CGG CAC TCG CGC GGT GGG CAC CCC TGG CTA TCA GGC CAG CGT CGA
arg gln arg gln arg arg his ser arg gly gly his pro trp leu ser gly gln arg arg
121/41                               151/51
CTA TGT GGT AAA CAC ACT GCG CAA CAG CGG TTT TGA TGT GCA AAC CCC GGA GTT CTC CGC
leu cys gly lys his thr ala gln gln arg phe opa cys ala asn pro gly val leu arg
181/61                               211/71
TCG CGT GTT CAA GGC CGA AAA AGG GGT GGT GAC CCT CGG CGG CAA CAC CGT CGA GGC GAG
ser arg val gln gly arg lys arg gly gly asp pro arg arg gln his arg gly gly glu
241/81                               271/91
GGC GCT CGA GTA CAG CCT CGG CAC ACC GCC GGA CGG GGT GAC GGG CCC GCT GGT GGC TGC
gly ala arg val gln pro arg his thr ala gly arg gly asp gly pro ala gly gly cys
301/101                               331/111
CCC CGC CGA CGA CAG TCC GGG CTG CAG TCC GTC GGA CTA CGA CAG GCT GCC GGT GTC CGG
pro arg arg arg gln ser gly leu gln ser val gly leu arg gln ala ala gly val arg
361/121
TGC GGT GGT GCT GGT AGA TC
cys gly gly ala gly arg

```

SEQ ID N° 19

FIGURE 23A

DE14.strider [2 to 380] -&gt; 1-phase Translation

DNA sequence 380 b.p. GTTGGCAACGG ... TGCTGGTAGATC linear

```

2/1                                32/11
TTG CGC AAC GGG GTG AGC ACC GAC GCG ATG ATG GCG CAA CTA TCG AAA CTG CAG GAC ATC
leu arg asn gly val ser thr asp ala met met ala gln leu ser lys leu gln asp ile
62/21                                92/31
GCC AAC GCC AAC GAC GGC ACT CGC GCG GTG GCG ACC CCT GGC TAT CAG GCC AGC GTC GAC
ala asn ala asn asp gly thr arg ala val gly thr pro gly tyr gln ala ser val asp
122/41                                152/51
TAT GTG GTA AAC ACA CTG CGC AAC AGC GGT TTT GAT GTG CAA ACC CCG GAG TTC TOC GCT
tyr val val asn thr leu arg asn ser gly phe asp val gln thr pro glu phe ser ala
182/61                                212/71
CGC GTG TTC AAG GCC GAA AAA GGG GTG GTG ACC CTC GGC GGC AAC ACC GTG GAG GCG AGG
arg val phe lys ala glu lys gly val val thr leu gly gly asn thr val glu ala arg
242/81                                272/91
GCG CTC GAG TAC AGC CTC GGC ACA CCG CCG GAC GGG GTG ACG GGC CCG CTG GTG GCT GCC
ala leu glu tyr ser leu gly thr pro pro asp gly val thr gly pro leu val ala ala
302/101                                332/111
CCC GCC GAC GAC AGT CCG GGC TGC AGT CCG TCG GAC TAC GAC AGG CTG CCG GTG TOC GGT
pro ala asp asp ser pro gly cys ser pro ser asp tyr asp arg leu pro val ser gly
362/121
GCG GTG GTG CTG GTA GAT C
ala val val leu val asp

```

SEQ ID N° 19

FIGURE 23B

## DE14.strider [3 to 380] -&gt; 1-phase Translation

DNA sequence 380 b.p. GTTGGCAACGG ... TGCTGGTAGATC linear

```

3/1
TGC GCA ACG GGG TGA GCA CCG ACG CGA TGA TGG CGC AAC TAT CGA AAC TGC AGG ACA TCG
cys ala thr gly OPA ala pro thr arg OPA trp arg asn tyr arg asn cys arg thr ser
63/21
CCA ACG CCA ACG ACG GCA CTC GCG CGG TGG GCA CCC CTG GCT ATC AGG CCA GCG TCG ACT
pro thr pro thr thr ala leu ala arg trp ala pro leu ala ile arg pro ala ser thr
123/41
ATG TGG TAA ACA CAC TGC GCA ACA GCG GTT TIG ATG TGC AAA CCC CGG AGT TCT CCG CTC
met trp OCH thr his cys ala thr ala val leu met cys lys pro arg ser ser pro leu
183/61
GCG TGT TCA ACG CCG AAA AAG GGG TGG TGA CCC TCG GCG GCA ACA CCG TGG AGG CGA GGG
ala cys ser arg pro lys lys gly trp OPA pro ser ala ala thr pro trp arg arg gly
243/81
CGC TCG AGT ACA GCC TCG GCA CAC CGC CGG ACG GGG TGA CGG GCC CGC TGG TGG CTG CCC
arg ser ser thr ala ser ala his arg arg thr gly OPA arg ala arg trp trp leu pro
303/101
CGG CCG ACG ACA GTC CGG GCT GCA GTC CGT CGG ACT ACG ACA GGC TGC CGG TGT CCG GTG
pro pro thr thr val arg ala ala val arg arg thr thr thr gly cys arg cys pro val
363/121
CGG TGG TGC TGG TAG ATC
arg trp cys trp AMB ile

```

SEQ ID N° 19

FIGURE 23C

## DE15.strider -&gt; 1-phase Translation

DNA sequence 390 b.p. CGAGACAGTGGT ... TGCATCAGGATC linear

```

1/1                               31/11
CGA GAC AGT GGT GCG GGA CAC TTG AGT TCG GCT GCT AAC GAC GCC AGA GTC GCC CGC TTC
arg asp ser gly ala gly his leu ser ser ala ala asn asp ala arg val ala arg phe
61/21                               91/31
CGC GGT GTG GGA CTC ACG TTC GGT GAG GGT ACA GCG GAC CTT CGA GCA CGC AAT ATC GTG
arg gly val gly leu thr phe gly glu gly thr ala asp leu arg ala arg asn ile val
121/41                             151/51
GGC CGG CTG GCA ACC GTC GGT TTC GAC GTT GGT GAC GAC CCC TCG TTC ATG AAT CGT TCT
gly arg leu ala thr val gly phe asp val gly asp asp pro ser phe met asn arg ser
181/61                             211/71
TGA GCT CCC CGT TTT GCT GGA TGC CCA GGC ACC GCC GGT ACT GCT GCG CTT AAG CTT GTC
OPA ala pro arg phe ala gly cys pro gly thr ala gly thr ala ala leu lys leu val
241/81                             271/91
GCA CAT GGT GGC GGC AGG GAG GAA CAG TGG GCA AGC AGC TAG CCG CGC TCG CCG CGC TGG
ala his gly ala gly arg glu glu gln trp ala ser ser AMB pro arg ser pro arg trp
301/101                           331/111
TCG GTG CGT GCA TGC TCG CAG CCG GAT GCA CCA ACG TCG TCG ACG GCA CCG CCG TGG CTG
ser val arg ala cys ser gln pro asp ala pro thr trp ser thr gly pro pro trp leu
361/121
CCG ACA AAT CCG GAC CAC TGC ATC AGG ATC
pro thr asn pro asp his cys ile arg ile

```

SEQ ID N° 20

FIGURE 24A



## DE15.strider [2 to 390] -&gt; 1-phase Translation

DNA sequence 390 b.p. CGAGACAGTGGT ... TGCATCAGGATC linear

2/1 32/11  
 GAG ACA GTG GTG CCG GAC ACT TGA GTT CCG CTG CTA ACG ACG CCA GAG TCG CCC GCT TCC  
 glu thr val val arg asp thr OPA val arg leu leu thr thr pro glu ser pro ala ser  
 62/21 92/31  
 GCG GTG TGG GAC TCA CGT TCG GTG AGG GTA CAG CCG ACC TTC GAG CAC GCA ATA TCG TGG  
 ala val trp asp ser arg ser val arg val gln arg thr phe glu his ala ile ser trp  
 122/41 152/51  
 GCC GGC TGG CAA CCG TCG GTT TCG ACG TTG GTG ACG ACC CCT CGT TCA TGA ATC GTT CTT  
 ala gly trp gln pro ser val ser thr leu val thr thr pro arg ser OPA ile val leu  
 182/61 212/71  
 GAG CTC CCC GTT TTG CTG GAT GCC CAG GCA CCG CCG GTA CTG CTG CGC TTA AGC TTG TCG  
 glu leu pro val leu leu asp ala gln ala pro pro val leu leu arg leu ser leu ser  
 242/81 272/91  
 CAC ATG GTG CCG GCA GGG AGG AAC AGT GGG CAA GCA GCT AGC CGC GCT CGC CGC GCT GGT  
 his met val pro ala gly arg asn ser gly gln ala ala ser arg ala arg arg ala gly  
 302/101 332/111  
 CGG TGC GTG CAT GCT CGC AGC CCG ATG CAC CAA CGT GGT CGA CCG GAC CGC CGT GCC TGC  
 arg cys val his ala arg ser arg met his gln arg gly arg arg asp arg arg gly cys  
 362/121  
 CGA CAA ATC CCG ACC ACT GCA TCA GGA TC  
 arg gln ile arg thr thr ala ser gly

SEQ ID N° 20

FIGURE 24B

## DE15.strider [3 to 390] -&gt; 1-phase Translation

DNA sequence 390 b.p. CGAGACAGTGGT ... TGCATCAGGATC linear

3/1 33/11  
 AGA CAG TGG TGC GGG ACA CTT GAG TTC GGC TGC TAA CGA CGC CAG AGT CGC CCG CTT CCG  
 arg gln trp cys gly thr leu glu phe gly cys OCH arg arg gln ser arg pro leu pro  
 63/21 93/31  
 CCG TGT GGG ACT CAC GTT CCG TGA GGG TAC AGC GGA CCT TCG ACC ACG CAA TAT CGT GGG  
 arg cys gly thr his val arg OPA gly tyr ser gly pro ser ser thr gln tyr arg gly  
 123/41 153/51  
 CCG GCT GGC AAC CGT CCG TTT CGA CGT TGG TGA CGA CCC CTC GTT CAT GAA TCG TTC TTG  
 pro ala gly asn arg arg phe arg arg trp OPA arg pro leu val his glu ser phe leu  
 183/61 213/71  
 AGC TCC CCG TTT TGC TGG ATG CCC AGG CAC CGC CGG TAC TGC TGC GCT TAA GCT TGT CGC  
 ser ser pro phe cys trp met pro arg his arg arg tyr cys cys ala OCH ala cys arg  
 243/81 273/91  
 ACA TGG TGC CCG CAG GGA GGA ACA GTG GGC AAG CAG CTA GCC GCG CTC GCC GCG CTG GTC  
 thr trp cys arg gln gly gly thr val gly lys gln leu ala ala leu ala ala leu val  
 303/101 333/111  
 GGT GCG TGC ATG CTC GCA GCC GGA TGC ACC AAC GTG GTC GAC GGG ACC GCC GTG GCT GCC  
 gly ala cys met leu ala ala gly cys thr asn val val asp gly thr ala val ala ala  
 363/121  
 GAC AAA TCC GGA CCA CTG CAT CAG GAT C  
 asp lys ser gly pro leu his gln asp

SEQ ID N° 20

FIGURE 24C

61/73

## DE16.strider -&gt; 1-phase Translation

DNA sequence 390 b.p. GTCTGGTGGCC ... GCTTGGGGATC linear

```

1/1                               31/11
GTC CTG GTC GGC GCG CAA CTG GCC GGT CCC GAT GGA AAG TGT TCA CGA TCG CGC TTC TGC
val leu val ala ala gln leu ala gly pro asp gly lys cys ser arg ser arg phe cys
61/21                               91/31
CGC TGG TAG TGG CGA TGG TGT TAG CAG GAT TGC GGG TCG AGG CTG CGA TGG CCA GCA CCA
arg trp AMB trp arg trp cys AMB gln asp cys gly ser arg leu arg trp pro ala pro
121/41                               151/51
GCG GCC TGC GGC TGG TCG CCG CGC GCG CCG AAA TGA TAC CCG CGA TCA CGA AAT ACA TGT
ala ala cys gly trp ser pro arg ala pro lys OPA tyr pro arg ser arg asn thr cys
181/61                               211/71
CGG CGC TGG ACG TCG CCG TGC TGG CCA GCT CGA CCG GAC ACG ATG TGG AGG GGG CGC AGA
arg arg trp thr ser pro cys trp pro ala arg pro asp thr met trp arg gly arg arg
241/81                               271/91
AAA ACT TCA CCG CCC GCA AGT ACG AGC TGC AGA CGC GAC TGG CCG ACA CCG ACG TCA TCG
lys thr ser pro pro ala ser thr ser cys arg arg asp trp pro thr pro thr ser ser
301/101                               331/111
CAG ACG TGC GGT CCG GAG TGA ACA CGC TGC TCA ACG GCG GTC AGG CGC TGC TGG ATA AGA
gln thr cys gly arg glu OPA thr arg cys ser thr ala val arg arg cys trp ile arg
361/121
TGC TGG CCG ACA GCA TCG GCT TGC GGG ATC
cys trp pro thr ala ser ala cys gly ile

```

SEQ ID N° 21

FIGURE 25A

## DE16.strider [2 to 390] -&gt; 1-phase Translation

DNA sequence 390 b.p. GTCTGGTGGCC ... GCTTGCGGGATC linear

2/1 32/11  
 TCC TGG TCG CCG CGC AAC TGG CCG GTC CCG ATG GAA AGT GTT CAC GAT CGC GCT TCT GCC  
 ser trp ser pro arg asn trp pro val pro met glu ser val his asp arg ala ser ala  
 62/21 92/31  
 GCT GGT AGT GGC GAT GGT GTT AGC AGG ATT GCG GGT CGA GGC TGC GAT GGC CAG CAC CAG  
 ala gly ser gly asp gly val ser arg ile ala gly arg gly cys asp gly gln his gln  
 122/41 152/51  
 CCG CCT GCG GCT GGT CGC CGC GCG CGC CGA AAT GAT ACC CGC GAT CAC GAA ATA CAT GTC  
 arg pro ala ala gly arg arg ala arg arg asn asp thr arg asp his glu ile his val  
 182/61 212/71  
 GGC GCT GGA CGT CGC CGT GCT GGC CAG CTC GAC CGG ACA CGA TGT GGA GGG GGC GCA GAA  
 gly ala gly arg arg arg ala gly gln leu asp arg thr arg cys gly gly gly ala glu  
 242/81 272/91  
 AAA CTT CAC CGC CCG CAA GTA CGA GCT GCA GAC GCG ACT GGC CGA CAC CGA CGT CAT CGC  
 lys leu his arg pro gln val arg ala ala asp ala thr gly arg his arg arg his arg  
 302/101 332/111  
 AGA CGT GCG GTC GGG AGT GAA CAC GCT GCT CAA CGG CGG TCA GGC GCT GCT GGA TAA GAT  
 arg arg ala val gly ser glu his ala ala gln arg arg ser gly ala ala gly OCH asp  
 362/121  
 GCT GGC CGA CAG CAT CGG CTT GCG GGA TC  
 ala gly arg gln his arg leu ala gly

SEQ ID N° 21

FIGURE 25B

## DE16.strider [3 to 390] -&gt; 1-phase Translation

DNA sequence 390 b.p. GTCCTGGTGGCC ... GCTTGCGGGATC linear

```

3/1                                33/11
CCT GGT CGC CGC GCA ACT GGC CGG TCC CGA TGG AAA GTG TTC ACG ATC GCG CTT CTG CCG
pro gly arg arg ala thr gly arg ser arg trp lys val phe thr ile ala leu leu pro
63/21                                93/31
CTG GTA GTG GCG ATG GTG TTA GCA GGA TTG CCG GTC GAG GCT GCG ATG GCC AGC ACC AGC
leu val val ala met val leu ala gly leu arg val glu ala ala met ala ser thr ser
123/41                                153/51
GGC CTG CCG CTG GTC GCC GCG CGC GCC GAA ATG ATA CCC GCG ATC ACG AAA TAC ATG TCG
gly leu arg leu val ala ala arg ala glu met ile pro ala ile thr lys tyr met ser
183/61                                213/71
GCG CTG GAC GTC GCC GTG CTG GCC AGC TCG ACC GGA CAC GAT GTG GAG GGG GCG CAG AAA
ala leu asp val ala val leu ala ser ser thr gly his asp val glu gly ala gln lys
243/81                                273/91
AAC TTC ACC GCC CGC AAG TAC GAG CTG CAG ACG CGA CTG GCC GAC ACC GAC GTC ATC GCA
asn phe thr ala arg lys tyr glu leu gln thr arg leu ala asp thr asp val ile ala
303/101                                333/111
GAC GTG CCG TCG GGA GTG AAC ACG CTG CTC AAC GGC GGT CAG GCG CTG CTG GAT AAG ATG
asp val arg ser gly val asn thr leu leu asn gly gly gln ala leu leu asp lys met
363/121
CTG GCC GAC AGC ATC GGC TTG CCG GAT C
leu ala asp ser ile gly leu arg asp

```

SEQ ID N° 21

FIGURE 25C

64/73

## DE21.strider -&gt; 1-phase Translation

DNA sequence 420 b.p. CTACGACAAGGC ... CACTACAAGATC linear

```

1/1                               31/11
CTA CGA CAA GGC AAA GGA GCA CAC GGT GAA GCG TGG ACT GAC GGT CGC GGT AGC CGG AGC
leu arg gln gly lys gly ala gln gly glu ala trp thr asp gly arg gly ser arg ser
61/21                               91/31
CGC CAT TCT GGT CGC AGG TCT TTC CGG ATG TTC AAG CAA CAA GTC GAC TAC AGG AAG CGG
arg his ser gly arg arg ser phe arg met phe lys gln gln val asp tyr arg lys arg
121/41                               151/51
TGA GAC CAC GAC CGC GGC AGG CAC GAC GGC AAG CCC CGG CGC CGC ATC CGG GCC GAA GGT
OPA asp his asp arg gly arg his asp gly lys pro arg arg arg ile arg ala glu gly
181/61                               211/71
CGT CAT CGA CGG TAA GGA CCA GAA CGT CAC CGG GTC TGT GGT GTG CAC AAC CGC GCC CGG
arg his arg arg OCH gly pro glu arg his arg val cys gly val his asn arg gly arg
241/81                               271/91
CAA TGT CAA CAT CGC GAT CGG CGG GGC GGC GAC CGG CAT TGC CGC CGT GCT CAC CGA CGG
gln cys gln his arg asp arg arg gly gly asp arg his cys arg arg ala his arg arg
301/101                               331/111
CAA CCC TCC GGA GGT GAA GTC CGT TGG GCT CGG TAA CGT CAA CGG CGT CAC GCT GGG ATA
gln pro ser gly gly glu val arg trp ala arg OCH arg gln arg arg his ala gly ile
361/121                               391/131
CAC GTC GGG CAC CGG ACA GGG TAA CGC TCG GCA ACC AAG GAC GGC AGC CAC TAC AAG ATC
his val gly his arg thr gly OCH arg ser ala thr lys asp gly ser his tyr lys ile

```

SEQ ID N° 22

FIGURE 26A

65/73

DE21.strider [2 to 420] -&gt; 1-phase Translation

DNA sequence 420 b.p. CTACGACAAGGC ... CACTACAAGATC linear

2/1 32/11 -  
 TAC GAC AAG GCA AAG GAG CAC AGG GTG AAG CGT GGA CTG ACG GTC GCG GTA GCC GGA GCC  
 tyr asp lys ala lys glu his arg val lys arg gly leu thr val ala val ala gly ala  
 62/21 92/31  
 GCC ATT CTG GTC GCA GGT CTT TCC GGA TGT TCA AGC AAC AAG TCG ACT ACA GGA AGC GGT  
 ala ile leu val ala gly leu ser gly cys ser ser asn lys ser thr thr gly ser gly  
 122/41 152/51  
 GAG ACC ACG ACC GCG GCA GGC ACG ACG GCA AGC CCG GGC GCC GCA TCC GGG CCG AAG GTC  
 glu thr thr thr ala ala gly thr thr ala ser pro gly ala ala ser gly pro lys val  
 182/61 212/71  
 GTC ATC GAC GGT AAG GAC CAG AAC GTC ACC GGG TCT GTG GTG TGC ACA ACC GCG GCC GGC  
 val ile asp gly lys asp gln asn val thr gly ser val val cys thr thr ala ala gly  
 242/81 272/91  
 AAT GTC AAC ATC GCG ATC GGC GGG GCG GCG ACC GGC ATT GCC GCC GTG CTC ACC GAC GGC  
 asn val asn ile ala ile gly gly ala ala thr gly ile ala ala val leu thr asp gly  
 302/101 332/111  
 AAC CCT CCG GAG GTG AAG TCC GTT GGG CTC GGT AAC GTC AAC GGC GTC ACG CTG GGA TAC  
 asn pro pro glu val lys ser val gly leu gly asn val asn gly val thr leu gly tyr  
 362/121 392/131  
 ACG TCG GGC ACC GGA CAG GGT AAC GCT CCG CAA CCA AGG ACG GCA GCC ACT ACA AGA TC  
 thr ser gly thr gly gln gly asn ala arg gln pro arg thr ala ala thr thr arg

SEQ ID N° 22

FIGURE 26B

DE21.strider [3 to 420] -> 1-phase Translation

DNA sequence 420 b.p. CTACGACAAGGC ... CACTACAAGATC linear

```

3/1                                33/11
ACG ACA AGG CAA AGG AGC ACA GGG TGA AGC GTG GAC TGA CGG TCG CGG TAG CCG GAG CCG
thr thr arg gln arg ser thr gly OPA ser val asp OPA arg ser arg AMB pro glu pro
63/21                                93/31
CCA TTC TGG TCG CAG GTC TTT CCG GAT GTT CAA GCA ACA AGT CGA CTA CAG GAA GCG GTG
pro phe trp ser gln val phe pro asp val gln ala thr ser arg leu gln glu ala val
123/41                                153/51
AGA CCA CGA CCG CCG CAG GCA CGA CCG CAA GCG CCG CCG CAT CCG GCG CGA AGC TCG
arg pro arg pro arg gln ala arg arg gln ala pro ala pro his pro gly arg arg ser
183/61                                213/71
TCA TCG ACC GTA AGG ACC AGA ACG TCA CCG GGT CTG TGG TGT GCA CAA CCG CCG CCG GCA
ser ser thr val arg thr arg thr ser pro gly leu trp cys ala gln pro arg pro ala
243/81                                273/91
ATG TCA ACA TCG CGA TCG GCG GGG CCG CGA CCG GCA TTG CCG CCG TGC TCA CCG ACG GCA
met ser thr ser arg ser ala gly arg arg pro ala leu pro pro cys ser pro thr ala
303/101                                333/111
ACC CTC CCG AGG TGA AGT CCG TTG GGC TCG GTA ACG TCA ACG GCG TCA CCG TGG GAT ACA
thr leu arg arg OPA ser pro leu gly ser val thr ser thr ala ser arg trp asp thr
363/121                                393/131
CGT CCG GCA CCG GAC ACG GTA ACG CTC GGC AAC CAA GGA CCG CAG CCA CTA CAA GAT C
arg arg ala pro asp arg val thr leu gly asn gln gly arg gln pro leu gln asp

```

SEQ ID N° 22

FIGURE 26C



## DE22.strider -&gt; 1-phase Translation

DNA sequence 597 b.p. GCACAACCGCGG ... CACTACAAGATC linear

```

1/1                                31/11
GCA CAA CCG CCG CCG GCA ATG TCA ACA TCG CGA TCG GCG GGG CCG CGA CCG GCA TTG CCG
ala gln pro arg pro ala met ser thr ser arg ser ala gly arg arg pro ala leu pro
61/21                                91/31
CCG TGC TCA CCG ACG GCA ACC CTC CCG AGG TGA AGT CCG TTG GCG TCG GTA ACC TCA ACC
pro cys ser pro thr ala thr leu arg arg OPA ser pro leu gly ser val thr ser thr
121/41                                151/51
GCG TCA CCG TGG GAT ACA CGT CCG GCA CCG GAC AGG GTA ACG CCT CCG CAA CCA ACG ACG
ala ser arg trp asp thr arg arg ala pro asp arg val thr pro arg gln pro arg thr
181/61                                211/71
GCA GCC ACT ACA AGA TCA CAG GGT GAA GCG TGG ACT GAC GGT CCG GGT ACG CCG ACG CCG
ala ala thr thr arg ser gln gly glu ala trp thr asp gly arg gly ser arg ser arg
241/81                                271/91
CAT TCT GGT CCG AGG TCT TTC CCG ATG TTC AAG CAA CAA GTC GAC TAC AGG AAG CCG TGA
his ser gly arg arg ser phe arg met phe lys gln gln val asp tyr arg lys arg OPA
301/101                                331/111
GAC CAC GAC CCG GGC AGG CAC GAC GGC AAG CCC CCG CCG CCC TOC GGG CCG AAG GTC GTC
asp his asp arg gly arg his asp gly lys pro arg arg arg ser gly pro lys val val
361/121                                391/131
ATC GAC GGT AAG GAC CAG AAC GTC ACC GGC TCC GTG GTG TGC ACA ACC GCG GCC GGC AAT
ile asp gly lys asp gln asn val thr gly ser val val cys thr thr ala ala gly asn
421/141                                451/151
GTC AAC ATC GCG ATC GGC GGG GCG GCG ACC GGC ATT GCC GCC GTG CTC ACC GAC GGC AAC
val asn ile ala ile gly gly ala ala thr gly ile ala ala val leu thr asp gly asn
481/161                                511/171
CCT CCG GAG GTG AAG TOC GTT GGG CTC GGT AAC GTC AAC GGC GTC ACG CTC GGA TAC ACG
pro pro glu val lys ser val gly leu gly asn val asn gly val thr leu gly tyr thr
541/181                                571/191
TCG GGC ACC GGA CAG GGT AAC GCC TCG GCA ACC AAG GAC GGC AGC CAC TAC AAG ATC
ser gly thr gly gln gly asn ala ser ala thr lys asp gly ser his tyr lys ile

```

SEQ ID N° 23

FIGURE 27A

DE22.strider [2 to 597] -&gt; 1-phase Translation

DNA sequence 597 b.p. GCACAACCGGG ... CACTACAAGATC linear

2/1 32/11  
 CAC AAC CGC GGC CGG CAA TGT CAA CAT CGC GAT CGG CGG GGC GGC GAC CGG CAT TCC CGC  
 his asn arg gly arg gln cys gln his arg asp arg arg gly gly asp arg his cys arg  
 62/21 92/31  
 CGT GCT CAC CGA OGG CAA OCC TCC GGA GGT GAA GTC CGT TGG GCT CGG TAA CGT CAA CGG  
 arg ala his arg arg gln pro ser gly gly glu val arg trp ala arg OCH arg gln arg  
 122/41 152/51  
 CGT CAC GCT GGG ATA CAC GTC GGG CAC CGG ACA GGG TAA CGC CTC GGC AAC CAA GGA CGG  
 arg his ala gly ile his val gly his arg thr gly OCH arg leu gly asn gln gly arg  
 182/61 212/71  
 CAG CCA CTA CAA GAT CAC AGG GTG AAG CGT GGA CTG ACG GTC GCG GTA GCC GGA GCC GCC  
 gln pro leu gln asp his arg val lys arg gly leu thr val ala val ala gly ala ala  
 242/81 272/91  
 ATT CTG GTC GCA GGT CTT TCC GGA TGT TCA AGC AAC AAG TCG ACT ACA GGA AGC GGT GAG  
 ile leu val ala gly leu ser gly cys ser ser asn lys ser thr thr gly ser gly glu  
 302/101 332/111  
 ACC ACG ACC GCG GCA GGC ACG ACG GCA AGC CCC GGC GGC GCT CCG GGC CGA AGG TCG TCA  
 thr thr thr ala ala gly thr thr ala ser pro gly ala ala pro gly arg arg ser ser  
 362/121 392/131  
 TCG ACG GTA AGG ACC AGA ACG TCA CCG GCT CCG TGG TGT GCA CAA CCG CGG CCG GCA ATG  
 ser thr val arg thr arg thr ser pro ala pro trp cys ala gln pro arg pro ala met  
 422/141 452/151  
 TCA ACA TCG CGA TCG GCG GGC CGG CGA CCG GCA TTG CCG CCG TGC TCA CCG ACG GCA ACC  
 ser thr ser arg ser ala gly arg arg pro ala leu pro pro cys ser pro thr ala thr  
 482/161 512/171  
 CTC CGG AGG TGA AGT CCG TTG GGC TCG GTA ACG TCA ACG GCG TCA CCG TGG GAT ACA CGT  
 leu arg arg OPA ser pro leu gly ser val thr ser thr ala ser arg trp asp thr arg  
 542/181 572/191  
 CGG GCA CCG GAC AGG GTA ACG CCT CGG CAA CCA AGG ACG GCA GCC ACT ACA AGA TC  
 arg ala pro asp arg val thr pro arg gln pro arg thr ala ala thr thr arg

SEQ ID N° 23

FIGURE 27B

## DE22.strider [3 to 592] -&gt; 1-phase Translation

DNA sequence 597 b.p. GCACAACCGCGG ... CACTACAAGATC linear

3/1  
 ACA ACC GCG GCC GGC AAT GTC AAC ATC GCG ATC GCG GGG GCG GCG ACC GGC ATT GCC GCC  
 thr thr ala ala gly asn val asn ile ala ile gly gly ala ala thr gly ile ala ala  
 63/21  
 GTG CTC ACC GAC GGC AAC OCT CCG GAG GTG AAG TCC GTT GGG CTC GGT AAC GTC AAC GGC  
 val leu thr asp gly asn pro pro glu val lys ser val gly leu gly asn val asn gly  
 123/41  
 GTC ACG CTG GGA TAC ACG TCG GGC ACC GGA CAG GGT AAC GCC TCG GCA ACC AAG GAC GGC  
 val thr leu gly tyr thr ser gly thr gly gln gly asn ala ser ala thr lys asp gly  
 183/61  
 AGC CAC TAC AAG ATC ACA GGG TGA AGC GTG GAC TGA CCG TCG CCG TAG CCG GAG CCG CCA  
 ser his tyr lys ile thr gly opa ser val asp opa arg ser arg amb pro glu pro pro  
 243/81  
 TTC TGG TCG CAG GTC TTT CCG GAT GTT CAA GCA ACA AGT CGA CTA CAG GAA GCG GTG AGA  
 phe trp ser gln val phe pro asp val gln ala thr ser arg leu gln glu ala val arg  
 303/101  
 CCA CGA CCG CCG CAG GCA CGA CCG CAA GCC CCG GCG CCG CTC CCG GCC GAA GGT CGT CAT  
 pro arg pro arg gln ala arg arg gln ala pro ala pro leu arg ala glu gly arg his  
 363/121  
 CGA CCG TAA GGA CCA GAA CGT CAC CCG CTC CGT GGT GTG CAC AAC CCG GGC CCG CAA TGT  
 arg arg och gly pro glu arg his arg leu arg gly val his asn arg gly arg gln cys  
 423/141  
 CAA CAT CCG GAT CCG CCG GCC GGC GAC CCG CAT TGC CCG CGT GCT CAC CGA CCG CAA CCC  
 gln his arg asp arg arg gly gly asp arg his cys arg arg ala his arg arg gln pro  
 483/161  
 TCC GGA GGT GAA GTC CGT TCG GCT CCG TAA CGT CAA CCG CGT CAC GCT GGG ATA CAC GTC  
 ser gly gly glu val arg trp ala arg och arg gln arg arg his ala gly ile his val  
 543/181  
 GGG CAC CCG ACA GGG TAA CCG CTC GGC AAC CAA GGA CCG CAG CCA CTA CAA GAT C  
 gly his arg thr gly och arg leu gly asn gln gly arg gln pro leu gln asp

SEQ ID N° 23

FIGURE 27C

## DE23.strider -&gt; 1-phase Translation

DNA sequence 420 b.p. CTAACGACAGGC ... CACTACAAGATC linear

```

1/1                               31/11
CTA ACG ACA GGC AAA GGA GCA CAG GGT GAA GCG TGG ACT GAC GGT CGC GGT AGC CGG AGC
leu thr thr gly lys gly ala gln gly glu ala trp thr asp gly arg gly ser arg ser
61/21
CGC CAT TCT GGT CGC AGG TCT TTC CGG ATG TTC AAG CAA CAA GTC GAC TAC AGG AAG CGG
arg his ser gly arg arg ser phe arg met phe lys gln gln val asp tyr arg lys arg
121/41
TGA GAC CAC GAC CGC GGC AGG CAC GAC GGC AAG CCC CGG CGC CGC TCC GGG CCG AAG GTC
OPA asp his asp arg gly arg his asp gly lys pro arg arg arg ser gly pro lys val
181/61
GTC ATC GAC GGT AAG GAC CAG AAC GTC ACC GGC TCC GTG GTG TGC ACA ACC CGC GGC GGC
val ile asp gly lys asp gln asn val thr gly ser val val cys thr thr ala ala gly
241/81
AAT GTC AAC ATC GCG ATC GGC GGG GCG GCG ACC GGC ATT GCC GCC GTG CTC ACC GAC GGC
asn val asn ile ala ile gly gly ala ala thr gly ile ala ala val leu thr asp gly
301/101
AAC CCT CCG GAG GTG AAG TCC GTT GGG CTC GGT AAC GTC AAC GCC GTC ACG CTG GGA TAC
asn pro pro glu val lys ser val gly leu gly asn val asn gly val thr leu gly tyr
361/121
ACG TCG GGC ACC GGA CAG GGT AAC GCC TCG GCA ACC AAG GAC GCC AGC CAC TAC AAG ATC
thr ser gly thr gly gln gly asn ala ser ala thr lys asp gly ser his tyr lys ile
391/131

```

SEQ ID N° 24

FIGURE 28A

71/73

DE23.strider [2 to 420] -&gt; 1-phase Translation

DNA sequence 420 b.p. CTAACGACAGGC ... CACTACAAGATC linear

2/1 32/11  
 TAA CGA CAG GCA AAG GAG CAC AGG GTG AAG CGT GGA CTG ACG GTC GCG GTA GCC GGA GCC  
 OCH arg gln ala lys glu his arg val lys arg gly leu thr val ala val ala gly ala  
 62/21 92/31  
 GCC ATT CTG GTC GCA GGT CTT TCC GGA TGT TCA AGC AAC AAG TCG ACT ACA GGA ACC GGT  
 ala ile leu val ala gly leu ser gly cys ser ser asn lys ser thr thr gly ser gly  
 122/41 152/51  
 GAG ACC ACG ACC GCG GCA GGC ACG ACG GCA AGC CCC GGC GCC GCT CCG GGC CGA AGG TCG  
 glu thr thr thr ala ala gly thr thr ala ser pro gly ala ala pro gly arg arg ser  
 182/61 212/71  
 TCA TCG ACG GTA AGG ACC AGA ACG TCA CCG GCT CCG TGG TGT GCA CAA CCG CCG CCG GCA  
 ser ser thr val arg thr arg thr ser pro ala pro trp cys ala gln pro arg pro ala  
 242/81 272/91  
 ATG TCA ACA TCG CGA TCG GCG GGG CCG CGA CCG GCA TTG CCG CCG TGC TCA CCG ACG GCA  
 met ser thr ser arg ser ala gly arg arg pro ala leu pro pro cys ser pro thr ala  
 302/101 332/111  
 ACC CTC CCG AGG TGA AGT CCG TTG GGC TCG GTA ACG TCA ACG GCG TCA CCG TGG GAT ACA  
 thr leu arg arg OPA ser pro leu gly ser val thr ser thr ala ser arg trp asp thr  
 362/121 392/131  
 CGT CCG GCA CCG GAC AGG GTA ACG CCT CCG CAA CCA AGG ACG GCA GCC ACT ACA AGA TC  
 arg arg ala pro asp arg val thr pro arg gln pro arg thr ala ala thr thr arg

SEQ ID N° 24

FIGURE 28B

72/73

DE23.strider [3 to 420] -&gt; 1-phase Translation

DNA sequence 420 b.p. CTAACGACAGGC ... CACTACAAGATC linear

```

3/1
AAC GAC AGG CAA AGG AGC ACA GGG TGA AGC GTG GAC TGA CGG TCG CGG TAG CCG GAG CCG
asn asp arg gln arg ser thr gly OPA ser val asp OPA arg ser arg AMB pro glu pro
63/21
CCA TTC TGG TCG CAG GTC TTT CCG GAT GTT CAA GCA ACA AGT CGA CTA CAG GAA GCG GTG
pro phe trp ser gln val phe pro asp val gln ala thr ser arg leu gln glu ala val
123/41
AGA CCA CGA CCG CGG CAG GCA CGA CGG CAA GCG CCG GCG CGC CTC CGG GCC GAA GGT CGT
arg pro arg pro arg gln ala arg arg gln ala pro ala pro leu arg ala glu gly arg
183/61
CAT CGA CCG TAA GGA CCA GAA CGT CAC CGG CTC CGT GGT GTG CAC AAC CGC GGC CGG CAA
his arg arg OCH gly pro glu arg his arg leu arg gly val his asn arg gly arg gln
243/81
TGT CAA CAT CGC GAT CGG CGG GGC GGC GAC CGG CAT TGC CGC CGT GCT CAC CGA CGG CAA
cys gln his arg asp arg arg gly gly asp arg his cys arg arg ala his arg arg gln
303/101
CCC TCC GGA GGT GAA GTC CGT TGG GCT CGG TAA CGT CAA CGG CGT CAC GCT GGG ATA CAC
pro ser gly gly glu val arg trp ala arg OCH arg gln arg arg his ala gly ile his
363/121
GTC GGG CAC CGG ACA GGG TAA CGC CTC GGC AAC CAA GGA CGG CAG CCA CTA CAA GAT C
val gly his arg thr gly OCH arg leu gly asn gln gly arg gln pro leu gln asp
393/131

```

SEQ ID N° 24

FIGURE 28C

SEQ ID N° 25

Amorce directe

5' ACG CGG CGC AGC CTG TTG 3'

SEQ ID N° 26

Amorce inverse

5' CGA CCT TGG GAT TCG CCT 3'

FIGURE 29

RAPPORT DE RECHERCHE  
PRELIMINAIREétabli sur la base des dernières revendications  
déposées avant le commencement de la rechercheN° d'enregistrement  
nationalFA 550795  
FR 9710404

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
X	WO 96 07745 A (PASTEUR INSTITUT ; GICQUEL BRIGITTE (FR); LIM ENG MONG (FR); PORTNO) 14 mars 1996 * abrégé * * page 1 - page 13 * * page 32 - page 37 * * page 45 - page 52; revendications * * figures 1,6 *	1-19,24, 40,42
A	SLAUCH J.M. AND SILHAVY T.J.: "Genetic fusions as experimental tools" METHODS IN ENZYMOLOGY, vol. 204, 1991, pages 213-248, XP002063253 * page 214, alinéa 2 - page 223, alinéa 1 *	1
A,D	ENG MONG LIM ET AL.: "Identification of Mycobacterium tuberculosis DNA sequences encoding exported proteins by using phoA gene fusions" JOURNAL OF BACTERIOLOGY, vol. 177, no. 1, 1 janvier 1995, pages 59-65, XP000560419	
A	TIMM J. ET AL.: "Escherichia coli-mycobacteria shuttle vectors for operon and gene fusions to lacZ: the pJEM series" JOURNAL OF BACTERIOLOGY, vol. 176, no. 21, novembre 1994, pages 6749-6753, XP002063184	
-/-		
Date d'achèvement de la recherche		Examineur
27 avril 1998		Macchia, G
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intermédiaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons &amp; : membre de la même famille, document correspondant</p>		

3

EPO FORM 1503 (3.92) (P4/C19)



INSTITUT NATIONAL  
de la  
PROPRIETE INDUSTRIELLE

**RAPPORT DE RECHERCHE  
PRELIMINAIRE**  
établi sur la base des dernières revendications  
déposées avant le commencement de la recherche

N° d'enregistrement  
national

FA 550795  
FR 9710404

DOCUMENTS CONSIDERES COMME PERTINENTS		Revendications concernées de la demande examinée
Catégorie	Citation du document avec indication, en cas de besoin, des parties pertinentes	
A,D	TIMM J. ET AL.: "Transcription and expression analysis, using lacZ and phoA gene fusions, of Mycobacterium fortuitum beta-lactamase genes cloned from a natural isolate and a high-level beta-lactamase producer" MOLECULAR MICROBIOLOGY, vol. 12, no. 3, 1994, pages 491-504, XP002063436	
A	DAS GUPTA S.K. ET AL.: "Cloning and assessment of mycobacterial promoters by using a plasmid shuttle vector" JOURNAL OF BACTERIOLOGY, vol. 175, no. 16, août 1993, pages 5186-5192, XP002063437	
		DOMAINES TECHNIQUES RECHERCHES (Int. CL. 6)
Date d'achèvement de la recherche		Examineur
27 avril 1998		Macchia, G
<p>CATEGORIE DES DOCUMENTS CITES</p> <p>X : particulièrement pertinent à lui seul Y : particulièrement pertinent en combinaison avec un autre document de la même catégorie A : pertinent à l'encontre d'au moins une revendication ou arrière-plan technologique général O : divulgation non-écrite P : document intercalaire</p> <p>T : théorie ou principe à la base de l'invention E : document de brevet bénéficiant d'une date antérieure à la date de dépôt et qui n'a été publié qu'à cette date de dépôt ou qu'à une date postérieure. D : cité dans la demande L : cité pour d'autres raisons &amp; : membre de la même famille, document correspondant</p>		

3

EPO FORM 1503 02.82 (P64C13)